Proceedings of the 7th International Conference on Inventive Computation Technologies (ICICT 2024)
Xplore Part Number : CFP24F70-ART ; ISBN : 979-8-3503-5929-9

2024 International Conference on Inventive Computation Technologies (ICICT) | 979-8-3503-5929-9/24/$31.00 ©2024 IEEE | DOI: 10.1109/ICICT60155.2024.10544592

Optimizing Well-Being: Unveiling Eustress and Distress through Machine Learning

Ashutosh Singh, Khushdeep Singh, Amit Kumar, Abhishek Shrivastava, and Santosh Kumar
ashutoshs22102@iiitnr.edu.in, khushdeep22102@iiitnr.edu.in, amit22102@iiitnr.edu.in,
abhisheks@iiitnr.edu.in and santosh@iiitnr.edu.in
Department of Data Science and Artificial Intelligence
IIIT Naya Raipur, Chhattisgarh, India 493661

***Abstract*—The predicament of rising stress levels among college students has increased rapidly, and its severe implications have caused numerous problems, such as notable anxiety, high blood pressure, increased cortisol levels, suicidal thoughts, and different health issues. This study addresses these issues found among college students by implementing machine learning algorithms like Decision Trees, Random Forest, Support Vector Machines, AdaBoost, Naive Bayes, Logistic Regression, and K-nearest Neighbors. The primary outcome of this work is to leverage a research study to predict and mitigate both the Eustress and Distress based on the collected questionnaire dataset. We conducted a workshop with the primary goal of studying the stress levels found among the students. This workshop was attended by students whose ages ranged between (18-21) years old, numbering about 800. A questionnaire was given to the students, validated under the guidance of the experts from All India Institute of Medical Sciences (AIIMS) in Raipur, Chhattisgarh, India, on which our dataset is based. The survey consists of 28 questions, aiming to comprehensively understand the multidimensional aspects of stress, including emotional well-being, physical health, academic performance, relationships, and leisure. The research finds that K-Nearest Neighbors has a maximum accuracy for Eustress, reaching 93%, and for Distress, the Logistic Regression has a maximum accuracy of 99%. The study contributes to a deeper understanding of stress determinants. It aims to improve college students' overall quality of life and academic success, addressing the multifaceted nature of stress.**

***Index Terms*—Stress, Eustress, Distress, Machine Learning, Logistic Regression, K-Nearest Neighbors.**

I. Introduction

The college experience for students is a transformative journey that challenges the mind and holds the power to shape character and foster ambition. This unique journey encourages self-exploration, providing personal and academic growth opportunities. Nevertheless, it is crucial to recognize that this journey is full of promise and potential and may also pose challenges that can significantly affect the well-being of students. Acknowledging their unique obstacles is essential in ensuring a supportive and empowering college experience tailored to their needs [15].

Eustress is a driving force that boosts performance, enables people to attain their goals, and gives them a sense of satisfaction. The thrill of studying, taking on challenging tasks, and participating in extracurricular activities can all be signs of Eustress throughout college. It is an essential element that fosters a positive outlook on life and personal development. Distress is the negative side of stress that is frequently associated with financial worries, personal conflicts, anxiety, depression, and academic pressure. Mental clarity, emotional stability, and physical health can all be negatively impacted by distress. As they negotiate demanding academic standards, social adaptations, and a variety of life transitions, college students frequently experience discomfort that can be difficult to manage [1].

A study [14] conducted by The Center of Healing (TCOH) in India reveals a surge in stress and anxiety levels among over 10,000 respondents since the onset of the COVID-19 pandemic [11]. The findings indicate that 74% of participants experience stress and 88% report anxiety. The study categorizes stress levels as 57% mild, 11% moderate, 4% moderately severe, and 2% severe [13].

The two distinct ways of identifying stress are defined as (1) physiological data and (2) psychological data. The major flaw in identifying stress by physiological data requires sensors to collect data under proper experimental conditions. This process can be both costly and time-consuming. In this work, we have used psychological data to identify Eustress and Distress because of its cost efficiency and easy implementation. We have used multiple machine learning algorithms to classify Eustress and Distress and validated them using various parameters such as accuracy, precision, recall, F1-score, and ROC curve to get the best algorithm for the stress dataset. These algorithms were tested and trained on the real-time dataset of college students collected by us.

II. Related Works

To increase our understanding and the quality of our work, we have investigated various jobs in this field by many researchers, as discussed in this section. Author [12] proposed a stress model for wrist-based electrodermal activity monitoring using machine learning techniques. The proposed framework is evaluated on different public stress datasets based on benchmark settings. The proposed framework provides 92.90% accuracy for detecting stress and the non-stress level of an individual. The major shortcoming of the proposed framework is that the proposed model might not capture minute changes in stress due to emotional changes in the individual.

Nath et al. [10] proposed a model using deep learning techniques in which a wristband-based stress detection framework was used using cortisol as a stress biomarker. They achieve a 94% accuracy. The major problem is the device's high cost and limited availability. Mahalakshmi et al. [2] studied 190 students, ages 14 to 18, and their calculated mental stress. They used several machine-learning models in which K-Nearest Neighbors achieved a higher accuracy of 87.2%. This study [3] proposed the EmpathicSchool model. This multimodal dataset provides a valuable resource for understanding stress through facial expressions and physiological signals, contributing to stress detection research and applications.

In [4], the author introduced the Tr-Estimate technique in a study specifically designed to diagnose Post-traumatic stress disorder (PTSD) in survivors at an early stage. This technology provides individualized care based on each patient's unique medical history while continuously monitoring the physiological signals. Sinha et al. [5] explore real-time stress prediction among students using KNN and Naive Bayes algorithms, with Naive Bayes demonstrating superior efficacy and potential for assisting university students and staff in swiftly determining stress levels within the campus premises. In this work, author [6] used the integration of IoT and machine learning. It is utilized to investigate workplace stress, with KNN achieving the highest accuracy. The author [7] has used the HINTS database to extract 26 variables. He has used artificial neural networks, random forests, gradient boosting, and logistics Regression to look into the factors of psychological distress and identify 20 variables. Barbayannis et al. [9] surveyed 843 college students to examine the correlation between academic stress levels and mental well-being. The study finds high academic stress correlates with poor mental well-being across genders and academic years, in which first-year students reported the most elevated stress.

A. Problem Statement

In the Previous studies, authors mainly focused on stress and non-stress. Still, for people in academics, especially students, it is essential to study what kind of stress they face, whether it is Eustress or Distress. College students face unique social barriers and socioeconomic issues that greatly influence their academic success and well-being. Despite increasing stress awareness in present-day educational settings, there is a significant research gap about the varied experiences of Eustress and Distress among college students. Observing the daily activities shows various impacts both Eustress and Distress have on college students. Early stress detection is crucial, as in the previous studies. At the same time, the sensor data can help with this. A significant issue is that putting sensors in different places is expensive and time-consuming, mainly because this is a large-scale survey.

B. Aim of this study

1) Research and implementation for Eustress and Distress.
2) Use machine learning to predict and intervene in Eustress and Distress.
3) Improve understanding of stress determinants for preventive measures.
4) Enhance student's quality of life and academic success by addressing stress.

III. Methodology

A. Proposed Architecture

The working of the proposed framework is illustrated in (Fig.1). The proposed framework is composed of several modules: (1) data preparation and collection, (2) data preprocessing, (3) feature selection, (4) splitting of the dataset, (5) creation of models, (6) model training and testing, (7) classification, and (8) result in Eustress and Distress. The detailed description of each module is given in the following subsections.

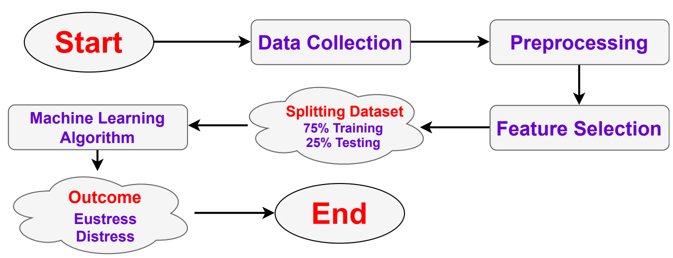

Fig. 1: Overview of the proposed method used for Eustress and Distress classification

B. Dataset description

We conducted a workshop to analyze students' stress levels with experts' help. In this workshop, different college students participated and were randomly given stress-related questionnaires. The real-time data was collected from various participating college students aged (18-21) years using a Google Form. The resulting dataset is organized in CSV format and encompasses 28 attributes for statistical analysis of stress levels provided by them.

These attributes correspond to a series of questions posed to college students, addressing their experiences and feelings concerning stress over the past two months. The questions have been categorized into seven distinct groups: Stress and Emotional Well-being, Physical Well-being, Academic Performance, Relationships and Social Environment, Leisure and Relaxation, and their Eustress and Distress levels. We have asked them for the necessary personal details. The questions have five response options from "Not at all" to "Extremely." These categories have been employed to label and organize the collected data shown in (Fig. 2), providing a comprehensive view of the students' stress-related experiences across these different dimensions [1] and [8].

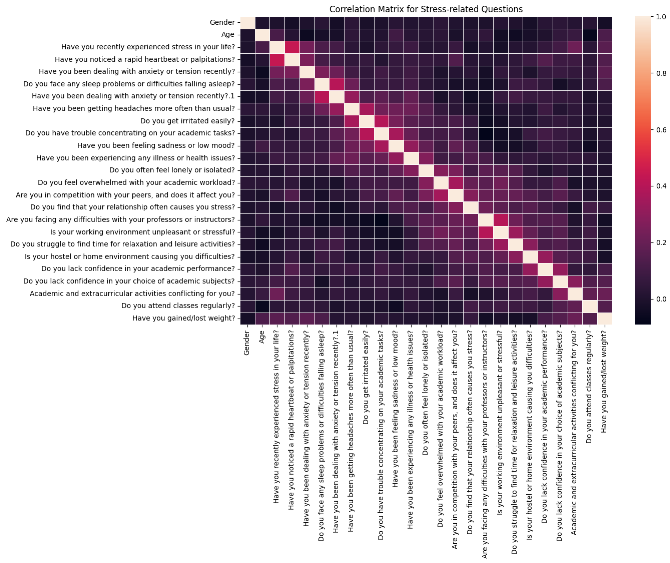

Fig. 2: Heatmap of the Eustress and Distress dataset

C. Preprocessing of collected dataset

Among the most critical issues during data processing are the techniques involving data preprocessing. This process encompasses various techniques to address missing values, anomalies, and inconsistencies. With the growing significance of data in research, business, and academia, more advanced analysis methods are required for larger datasets. For our model, we have used a few of the existing data preprocessing techniques, namely duplicate handling, encoding, and normalization, as shown in (Fig. 3).

In this work, the first step in the model is to collect the data. After data collection, there are many duplicate records in the data that have to be handled. Then, the process of normalization takes place. Then, we finally get the finalized normalized college dataset.

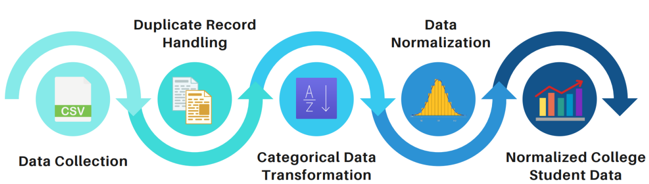

Fig. 3: Data Preprocessing overview for Eustress and Distress dataset

D. Algorithms

This study employed various machine learning algorithms to classify Eustress and Distress. Our findings indicate that Logistic Regression achieved the highest accuracy for Eustress, while the K-Nearest Neighbors model demonstrated superior performance for Distress.

1) Logistic Regression: Logistic Regression, a regression analysis method, aims to perform binary classification by modeling the probability of the default class. In our study, the primary aim was to utilize Logistic Regression for predicting stress severity. The algorithm aims to capture the relationship between input features and the probability of an instance belonging to a specific class, contributing to accurately classifying stress levels.

$$P(Y=1|\mathbf{x}) = \frac{1}{1+e^{-(\beta_0+\beta_1\cdot\mathbf{x})}} \tag{1}$$

2) k-Nearest Neighbors (k-NN): The k-Nearest Neighbors algorithm aims to classify instances based on their similarity to neighboring data points. It determines the majority class among the k-nearest neighbors of an example to make predictions. In our study, k-NN was employed for stress severity prediction, where the algorithm relies on the support of neighboring instances to determine the stress classification.

$$\hat{y} = \arg\max\left(\sum_{i=1}^{k} I(y_i = c)\right) \tag{2}$$

E. Statistical Measures

To find the statistical measures for accurate analysis of stress conditions, we used the statistical measure to evaluate the effectiveness of the proposed model. The statistical measures are depicted as follows.

$$\text{Precision} = \frac{TP}{TP+FP} \tag{3}$$

$$\text{Recall} = \frac{TP}{TP+FN} \tag{4}$$

$$\text{F1-score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision + Recall}} \tag{5}$$

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \tag{6}$$

IV. Results and Discussion

A. Performance Analysis

The performance of the proposed framework is evaluated based on 5-cross-validation benchmark settings. (Fig. 4 and Fig. 5) depict the overall performance analysis for the classification of stress severity for statistical measures. The results showed a comprehensive overview of the classification results for Eustress and Distress by utilizing the confusion matrix and ROC curve. The defined confusion matrix and ROC curve show the accurate classification for stress measurement. It is highly beneficial for studying the significant difference between Eustress and Distress, underscoring the efficacy of the classification model.

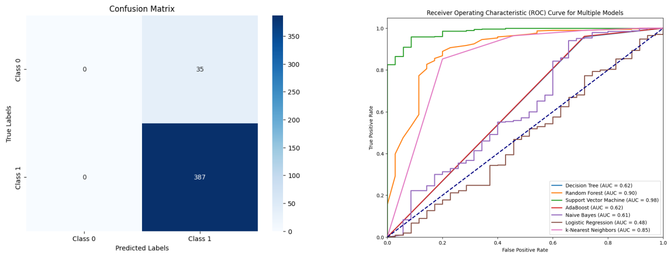

Fig. 4: (Left) confusion matrix and (Right) ROC curve of Eustress

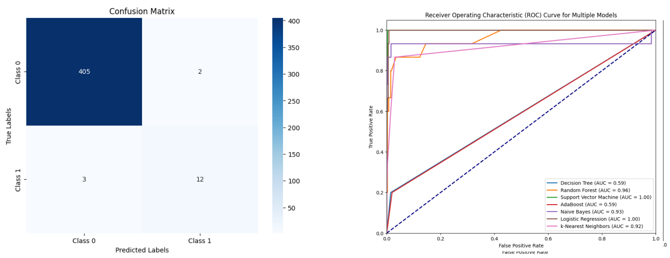

Fig. 5: (Left) confusion matrix and (Right) ROC curve of Distress

B. Comparison between Models

We compared the performance of the proposed system with state-of-the-art methods based on different settings, including several machine learning algorithms like Decision Trees (DT), Random Forest (RF), Support Vector Machines (SVM), Adaptive Boosting (AdaBoost), Naive Bayes (NV), Logistic Regression (LR), and K-nearest Neighbors (k-NN) for classify the levels of stress based on collected responses. (Fig. 6 and Fig. 7) shows the accuracy, precision, recall, and f1 score of the different types of stress classification. Based on overall observations, the stress classification using the existing methods is shown (Table. I and Table. II). Based on comprehensive analysis, the proposed framework provides a higher accuracy for classifying stress severity. It is highly used for the futuristic application for college students to better study their sentiments for selecting prospective opportunities in college activities.

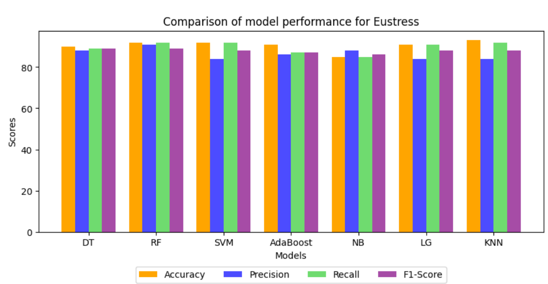

Fig. 6: comparing several machine learning models for Eustress

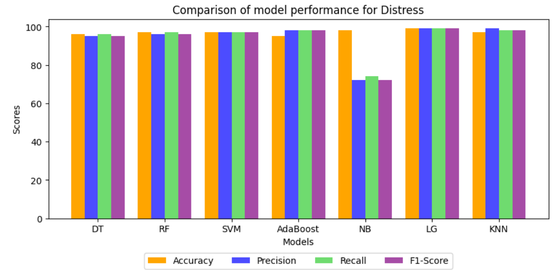

Fig. 7: Comparing several machine learning models for Distress

TABLE I: Comparison of model performance for Eustress

Model	Accuracy	Precision	Recall	F1-Score
DT	90	88	89	89
RF	92	91	92	89
SVM	92	84	92	88
AdaBoost	91	86	87	87
NB	85	88	85	86
LG	91	84	91	88
KNN	93	84	92	88

TABLE II: Comparison of model performance for Distress

Model	Accuracy	Precision	Recall	F1-Score
DT	96	95	96	95
RF	97	96	97	96
SVM	97	97	97	97
AdaBoost	95	98	98	98
NB	98	72	74	72
LG	99	99	99	99
KNN	97	99	98	98

V. Conclusion and Future Direction

In this work, we proposed a novel stress detection model using machine learning techniques to examine the predictive analysis of Eustress and Distress among college students based on prepared data. We evaluated the performance of different models based on different benchmark settings. Our results depicted that Logistic Regression techniques provided the most robust predictive accuracy among deployed models for stress assessment based on the prepared dataset, and it achieved 91% accuracy in classifying Eustress and a commendable 94% accuracy in classifying Distress. In contrast, K-Nearest Neighbors had the highest accuracy of 93% for Eustress and 97% for Distress. These findings underscore the potential utility of K-Nearest Neighbors and Logistic Regression as potent tools for proactively addressing the well-being of college students and facilitating early intervention where necessary. In the future, we will focus on improving the size of the collected stress dataset for analysis of different stress conditions. Moreover, we will prepare multimodal sensor data using wrist-worn wearable sensor devices for stress detection and classification of physiological activities of individuals using deep learning techniques.

REFERENCES

[1] Awada, Mohamad, Burcin Becerik-Gerber, Gale Lucas, Shawn Roll, and Ruying Liu. ”A New Perspective on Stress Detection: An Automated Approach for Detecting Eustress and Distress.” IEEE Transactions on Affective Computing (2023).

[2] Mahalakshmi, G., and G. Sujatha. ”Predictions of College Students’ Mental Stress using Machine Learning Algorithms.” In 2023 7th International Conference on Intelligent Computing and Control Systems (ICICCS), pp. 288-295. IEEE, (2023).

[3] Hosseini, Majid, Fahad Sohrab, Raju Gottumukkala, Ravi Teja Bhupatiraju, Satya Katragadda, Jenni Raitoharju, Alexandros Iosifidis, and Moncef Gabbouj. ”EmpathicSchool: A multimodal dataset for real-time facial expressions and physiological data analysis under different stress conditions.” arXiv preprint arXiv:2209.13542 (2022).

[4] Rachakonda, Laavanya, and K. C. Bipin. ”Tr-Estimate: A Novel Machine Learning Based Early Prediction System for Post-Traumatic Stress Disorder using IoMT.” In 2022 IEEE International Symposium on Smart Electronic Systems (iSES), pp. 677-682. IEEE, (2022).

[5] Sinha, Somnath, and R. Sriram. ”An Educational based Intelligent Student Stress Prediction using ML.” In 2022 3rd International Conference for Emerging Technology (INCET), pp. 1-7. IEEE, (2022).

[6] KS, SUHAS, and PHANEENDRA HD. ”Stress Prediction in Working Employees using Artificial Intelligence of Things.” Journal of Pharmaceutical Negative Results (2022).

[7] Chen, Yang, Xiaomei Zhang, Lin Lu, Yinzhi Wang, Jiajia Liu, Lei Qin, Linglong Ye, Jianping Zhu, Ben-Chang Shia, and Ming-Chih Chen. ”Machine learning methods to identify predictors of psychological distress.” Processes 10, no.5 (2022).

[8] Bisht, Aanchal, Shreya Vashisth, Muskan Gupta, and Ena Jain. ”Stress Prediction in Indian School Students Using Machine Learning.” In 2022 3rd International Conference on Intelligent Engineering and Management (ICIEM), pp. 770-774. IEEE, (2022).

[9] Barbayannis, Georgia, Mahindra Bandari, Xiang Zheng, Humberto Baquerizo, Keith W. Pecor, and Xue Ming. ”Academic stress and mental well-being in college students: correlations, affected groups, and COVID-19.” Frontiers in Psychology 13 (2022).

[10] Nath, Rajdeep Kumar, and Himanshu Thapliyal. ”Smart wristband-based stress detection framework for older adults with cortisol as stress biomarker.” IEEE Transactions on Consumer Electronics 67, no.1 (2021).

[11] Kumar, Santosh, Mithilesh Kumar Chaube, Saeed Hamood Alsamhi, Sachin Kumar Gupta, Mohsen Guizani, Raffaele Gravina, and Giancarlo Fortino. ”A novel multimodal fusion framework for early diagnosis and accurate classification of COVID-19 patients using X-ray images and speech signal processing techniques.” Computer methods and programs in biomedicine 226 (2022)

[12] Zhu, Lili, Petros Spachos, Pai Chet Ng, Yuanhao Yu, Yang Wang, Konstantinos Plataniotis, and Dimitrios Hatzinakos. ”Stress Detection Through Wrist-Based Electrodermal Activity Monitoring and Machine Learning.” IEEE Journal of Biomedical and Health Informatics (2023).

[13] Press Information Bureau ”Honouring and Empowering the Adivasis of India” https://encr.pw/r9oO7

[14] The Indian Express ”74% Indians suffering from stress, 88% from anxiety: Study” https://l1nq.com/vHG3K

[15] Stress in College Students: ”What to Know - U.S. News & World Report” https://shorturl.at/dNX07