Predicting the stress level of students using Supervised Machine Learning and Artificial Neural Network (ANN)

Article *in* Indian Journal of Engineering · August 2024

Nor Azuana Ramli
Universiti Malaysia Pahang Al-Sultan Abdullah

Indian journal of
Engineering

Predicting the stress level of students using Supervised Machine Learning and Artificial Neural Network (ANN)

Suraj Arya[1], **Anju**[2*], **Nor Azuana Ramli**[3]

To Cite:
Arya S, Anju, Ramli NA. Predicting the Stress level of students using Supervised Machine Learning and Artificial Neural Network (ANN). *Indian Journal of Engineering*, 2024, 21, e9ije1684
doi:

Author Affiliation:
[1]Assistant Professor, Department of Computer Science and Information Technology, Central University of Haryana, Mahendergarh, India
[2]Research Scholar, Department of Computer Science & Information Technology, Central University of Haryana, Mahendergarh, India
[3]Senior Lecturer, Center for Mathematical Sciences, University Malaysia Pahang AI-Sultan Abdullah, 26300, Kuantan, Pahang, Malaysia, India

***Corresponding Author**
Research Scholar, Department of Computer Science & Information Technology, Central University of Haryana,
Mahendergarh, India
Email: anju24sanga@gmail.com

Peer-Review History
Received: 02 May 2024
Reviewed & Revised: 06/May/2024 to 22/July/2024
Accepted: 26 July 2024
Published: 3 August 2024

Peer-Review Model
External peer-review was done through double-blind method.

Indian Journal of Engineering
pISSN 2319–7757; eISSN 2319–7765

DISCOVERY
SCIENTIFIC SOCIETY

ABSTRACT

Nowadays, the concept of stress is universally acknowledged. Many of us face situations that contribute to daily hassles, affecting professionals such as teachers, doctors, lawyers, journalists, and parents. University students are also encountering similar challenges. This study aims to identify the factors generating stress among students at Tribhuvan University Dharan in Nepal. We can predict and prevent stress at its early stages by analyzing these stress factors. This paper proposes various machine learning and deep learning models, including support vector machine (SVM), Random Forest, Gradient Boosting, AdaBoost, CatBoost, LightGBM, ExtraTree, XGBoost, logistic regression, K-nearest neighbor (KNN), Naive Bayes, decision tree, multi-layer perceptron (MLP), and artificial neural network (ANN). The Naive Bayes model achieved an accuracy of 90%, while SVM had the lowest test accuracy at 85.45%. The accuracy of these models improved with hyperparameter tuning. The key finding of this study is that the "academic period" is the most stressful time for students compared to other situations.

Keywords: Stress Prediction, Machine Learning, Random Forest, Naïve Bayes, Support Vector Machine, Artificial Neural Network.

1. INTRODUCTION

Stress is a state of mind in which a person feels pressured to perform daily routine activities. This phenomenon is evident across various sectors, including but not limited to offices, universities, hospitals, and others. Sometimes, it is obvious, but generally, it results from higher expectations and low passion, unrealistic workloads, insecure jobs, community violence, and examinations. Professionals like teachers, doctors, lawyers, journalists, parents, and others go through stressful situations. Even students are not spared from the stress. Students are the future of every country, so it is essential to analyze the factors responsible for stress among students. Thus, by earlier detection of these factors, stressful situations can be ignored or controlled. According to the World Health

Organization (WHO), stress is defined as a state of everyday pressure caused by a difficult situation (American Psychological Association, 2023).

Stress is classified into two types: Acute and chronic (Mind, 2022). Short-term stress is called acute stress. It can be seen in the examination hall, joining a new job, during speech, and facing a deadline for work. Chronic stress occurs when we feel pressure and tension for an extended period. It arises from financial difficulties, career-related problems, highly pressured jobs, and relationship challenges. It is essential to understand these factors for effective stress level management (Medlineplus, 2022). Stress management strategies can be developed with the help of machine learning and deep learning techniques. Physiological, psychological, academic, environmental, and social (PPAES) are the stress's various building blocks, as shown in Figure 1.

Stress factors can be integrated with machine learning (ML) and deep learning (DL) models to forecast student stress. In this paper, various ML models, such as support vector machines, decision trees, Random Forests, Extratree, and Naïve Bayes, are trained and tested. Stress levels are classified using labeled data. Applying hyperparameter tuning techniques has substantially improved the performance of machine learning (ML) algorithms. These algorithms are assessed using recall, precision, F1-score, and accuracy metrics.

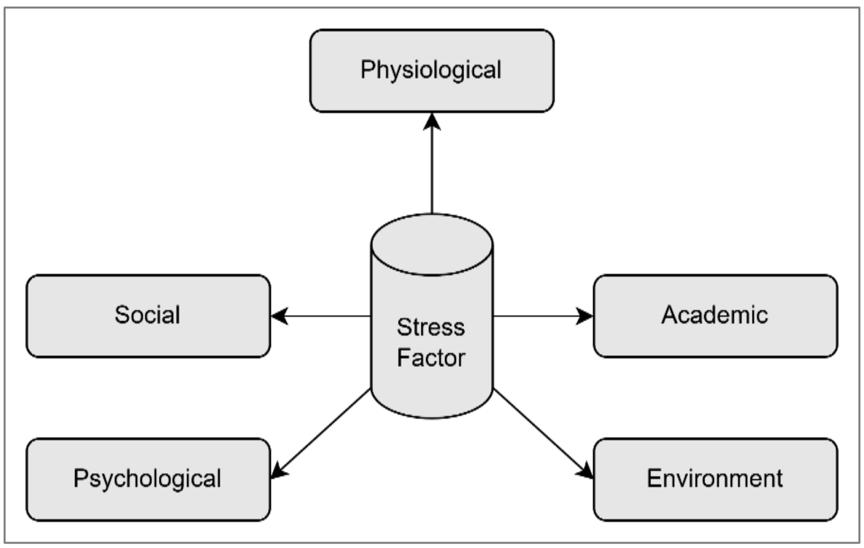

Figure 1 Stress factor classification of stress dataset

This paper is structured into five sections, where the second and third sections cover the related work and research methodology. At the same time, the experimental evaluation, results and discussions, and conclusion are detailed in the fourth and fifth sections, respectively.

Related Works

Stress expectation frameworks have been studied by several researchers using different machine-learning approaches. The researchers made numerous studies attempts to achieve effective methods and high accuracy in identifying stress-related elements. In their literature, they studied about three and four machine learning models, and others applied deep learning models. The existing works have been discussed here: Onim et al., (2024) used a fully connected convolutional neural network (CNN) model, which validates the integration of digital stress biomarkers. In this paper, they used a sensor with fusion and found 96.7525% accuracy and 0.9745 F1-score with the context in the case of CNN and Random Forest, providing accuracy and F1-score of 0.937 and 92.48%, respectively. In their investigation, Liao et al., (2024) examined elements in SRQ and SCL-90 and utilized machine learning to understand mental stress.

The SHAP model was used for finding the training results. The ROC-AUC curve has been evaluated for Random Forest. Ratul et al., (2023) examined the psychological and social levels of stress among university students. This type of stress was due to physical illness, mobile dependency, the internet, lack of social activities, and others. It examined 444 university students from various backgrounds. Different machine learning algorithms were used for feature reduction techniques: Principal component analysis (PCA) and the Chi-

Squared test. GridSearchCV, cross-validation, and GA were used to optimize hyperparameters. The result gave an accuracy of 80.5%, recall of 0.826, and precision of 1.00. Ding et al., (2023) were using machine learning and deep learning approaches for stress prediction. At present, stress increases health issues and puts human lives and health at risk, and 40% of young were facing stress due to frustration, nervousness, and anxiety. It was using a hybrid approach.

This consists of gradient boosting and Random Forest, which were given an accuracy of 100%, and a 10-fold cross-validation, also used for finding mean and standard deviation. A statistical T-test was applied to determine the relevance of this method in comparison with other machine learning methods. Al-Atawi et al., (2023) Employed machine learning, the Internet of Things (IoT), and wearable devices for managing stress. The stress was assessed using three main factors, such as body temperature, sweat, and movement in exercise, to achieve an accuracy of 99.5% and improve people's psychological health and well-being. "Humidity–Temperature–Step count–Stress levels" refers to a dataset with 2001 samples called Stress-Lysis.csv. This paper used machine learning techniques like stacked ensemble methods (SEM), gradient boosting, and Random Forest. Various sensors (temperature, humidity, and accelerometer) were used to measure stress levels.

Mittal et al., (2022) Focused on how stress is managed in the workplace and education field. It finds all feasible factors that were responsible for anxiety and depression. Various machine learning approaches, such as supervised and unsupervised techniques, have been used in this paper to detect stress effectively and efficiently among a vast population. It aims to detect stress before it occurs in life. It has also been used in deep learning and stress management, specifically in workplace and education settings. In a study conducted by Vallone et al., (2022), student stress levels were examined at four universities in Spain by surveying 331 students during the COVID-19 pandemic. The survey took place from 12 to 19 April 2021. Stress in students is due to institutional life, relationships, contagion, fear, and isolation. For this Coronavirus disease, the 2019 student stress questionnaire (CSSQ) had seven -items, and the original tool of the three-factor solution was used.

The Microsoft team-hosted online cross-sectional survey was used for this study. The objective of this CSSQ was to test structural validity, Discriminant validity, convergent validity, and internal consistency. According to Nijhawan et al., (2022), natural language processing (NLP) and ML have detected stress over social interaction. The latent Dirichlet Allocation (LDA) technique has been used for exploratory analysis of user tweets. A deep learning model, i.e., BERT, was used for sentimental classification, which had the best stress detection rate. According to Edwards et al., (2020), student stress has been reduced in the academic library using robot animal companions. After contact with a robot animal, participants reported decreased stress and increased positive affect.

One hundred three students from Midwestern US universities were selected for this study. Some libraries allow students to interact with pets such as dogs and cats. The authors organized a robot petting zoo at Midwestern University during the final week. Participants were asked to rate their stress level before interacting with T1, i.e., robot pet. After that interaction, participants were asked to assess their present state of stress and how supportive they thought the robot animal was (T2). Pairwise t-tests were employed to analyze the pre-test and post-test scores data. Table 1 shows the summary of existing studies.

Table 1 Literature Review of Existing Works

References	Methods/Techniques	Performance	Dataset/Survey/Link
Onim et al., (2024)	CNN, RF	CNN=96.7525% accuracy and 0.9745 F1 -score, RF= accuracy 0.937 and F1-score 92.48%	40 peoples of age between 60 to 80
Liao et al., (2024)	SRQ and SCL-90	SHAP, ROC AUC	https://studentlife.cs.dartmouth.edu/
Rescio et al., (2024)	1D-CNN, LSTM, and GRU	1D-CNN accuracy=95.38%	20 peoples (9 males, 11 females), average age=29.1, dataset is not publicly available
Mohamed et al., (2024)	DT, KNN, RF, SVC	KNN accuracy=98%	SWEET dataset of 1002 volunteers

Richer et al., (2024)	TSST and f-TSST	Accuracy=75±17.7 % and 73.4± 7.7 for the pilot and main study	Dataset available at https://osf.io/qvzdg/, https://osf.io/va6t3/ (59 participants)
Miao et al., (2023)	MFBPST-3D-DRLF	Accuracy of SEED = 96.67% and accuracy of SEED-IV=88.21%	SEED & SEED-IV dataset
Tang et al., (2023)	STILN	Arousal accuracy of =68.31% and valence accuracy=67.52%	Public DEAP dataset
Ratul et al., (2023)	Principle Component Analysis, Chi-Squared, Genetic Algorithms	accuracy (80.5%), precision (1.000), F1-score (0.890), and recall value (0.826)	Google Form of 444 students of university from different backgrounds/ ethnicities
Ding et al., (2023)	Hybrid Model (GB + RF), T-test	Mean accuracy (1), SD (+/-0.00)	Stress detection prediction
Al-Atawi et al., (2023)	RF, GB, Stacked Ensemble Method (SEM)	Accuracy (99.5%)	Stress-Lysis
Mittal et al., (2022)	SVM, DT, Neural Network, LDA, NB	SVM (99%) accuracy	College students Stress management data, Most common causes of stress at work in the United Kingdom in 2020
Vallone et al., (2022)	SPSS version 21 and AMOS tool version 26	-	Survey form (hosted by Microsoft Teams)
Nijhawan et al., (2022)	NLP, ML, BERT, LDA	-	100042 tweets had been used for binary sentimental
Edwards et al., (2020)	Academic Library by using robot animal	-	103 students from Midwestern US universities

The dataset from Tribhuvan University Dharan, Nepal is the original contribution and unique aspect of the study. This dataset has not previously been analyzed using ML and DL models to predict the stress factors among students. A total of twelve machine learning and two deep learning models were trained and tested for stress forecasting among students. Multiclass classification was used to forecast the stress level. Comparative analysis was performed to find out the most accurate model. Lastly, the stress level feature is not a previous analysis as a target variable by any researcher.

2. METHODOLOGY

This paper's data source is Kaggle, an open data repository. The students' stress dataset belongs to Dharan University of Nepal. After conducting the exploration analysis, we will conduct a comprehensive analysis of the stress_level feature. Several different criteria are used to compare the models under consideration. The results of all models are studied, and the best model is chosen. The GridSearchCV method is applied for hyperparameter adjustment to evaluate model performance relative to accuracy.

Dataset pre-processing

In the initial preprocessing stages, the first step involves importing the necessary library and reading the dataset. Subsequently, data cleaning is performed to identify and remove duplicate, irrelevant, and missing values. Any identified missing or duplicate values are then eliminated. The dataset pertains to Dharan Tribhuvan University, Nepal, and contains zero null values, indicating its readiness for direct utilization. As depicted in Figure 2, all columns exhibit uniform data type objects, an equal count, the absence of duplicate rows, and no missing values, thus confirming the dataset's integrity. Consequently, no further cleaning is deemed necessary.

As shown in Figure 3, we will divide the dataset into two parts: 80% for the training set and 20% for the testing set. Through the use of cross-validation and hyperparameter techniques in machine learning, the goal is to enhance the accuracy of the models and determine the optimal parameters for these algorithms.

```
anxiety_level                  0
self_esteem                    0
mental_health_history          0
depression                     0
headache                       0
blood_pressure                 0
sleep_quality                  0
breathing_problem              0
noise_level                    0
living_conditions              0
safety                         0
basic_needs                    0
academic_performance           0
study_load                     0
teacher_student_relationship   0
future_career_concerns         0
social_support                 0
peer_pressure                  0
extracurricular_activities     0
bullying                       0
stress_level                   0
dtype: int64
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1100 entries, 0 to 1099
Data columns (total 21 columns):
 #   Column                        Non-Null Count  Dtype
---  ------                        --------------  -----
 0   anxiety_level                 1100 non-null   int64
 1   self_esteem                   1100 non-null   int64
 2   mental_health_history         1100 non-null   int64
 3   depression                    1100 non-null   int64
 4   headache                      1100 non-null   int64
 5   blood_pressure                1100 non-null   int64
 6   sleep_quality                 1100 non-null   int64
 7   breathing_problem             1100 non-null   int64
 8   noise_level                   1100 non-null   int64
 9   living_conditions             1100 non-null   int64
 10  safety                        1100 non-null   int64
 11  basic_needs                   1100 non-null   int64
 12  academic_performance          1100 non-null   int64
 13  study_load                    1100 non-null   int64
 14  teacher_student_relationship  1100 non-null   int64
 15  future_career_concerns        1100 non-null   int64
 16  social_support                1100 non-null   int64
 17  peer_pressure                 1100 non-null   int64
 18  extracurricular_activities    1100 non-null   int64
 19  bullying                      1100 non-null   int64
 20  stress_level                  1100 non-null   int64
dtypes: int64(21)
memory usage: 180.6 KB
```

Figure 2 Identifying null, duplicates, and missing values

Model training process

Machine learning (ML) and deep learning (DL) models are instructed after pre-processing and normalizing the datasets. The dataset is randomly divided into training and testing data (80% training and 20% testing). The k-fold cross-validation and hyperparameter tuning are used to determine the accuracy metrics. Deep learning and machine learning techniques used are artificial neural network (ANN), multi-layer perceptron (MLP), support vector machine (SVM), Random Forest (RF), GradientBoosting (GB), AdaBoost, CatBoost, LightGBM, ExtraTree, XGBoost, logistic regression (LR), K-Nearest Neighbour (KNN), Naïve Bayes (NB), and decision tree (DT). These models are the most suitable for our research problem as they have features like regression and classification. ML and DL models used in this research work are as follows:

Supervised machine learning algorithms, such as SVM, can be used for classification and regression. SVM can find a hyperplane that separates maximal data into different classes. SVC uses a linear classifier with a straight-bit capacity. Linear SVCs have more restrictions, like the standardization of consequences and misfortune (Zhang et al., 2013). The machine learning algorithm, Random Forest, was coined by Leo Breiman and Adele Cutler, combining the results of multiple decision trees to arrive at one result. The adoption of this tool has been fueled by its ease of use and flexibility as it overcomes classification and regression problems. This algorithm solves the overfitting of the decision tree (Liu et al., 2012).

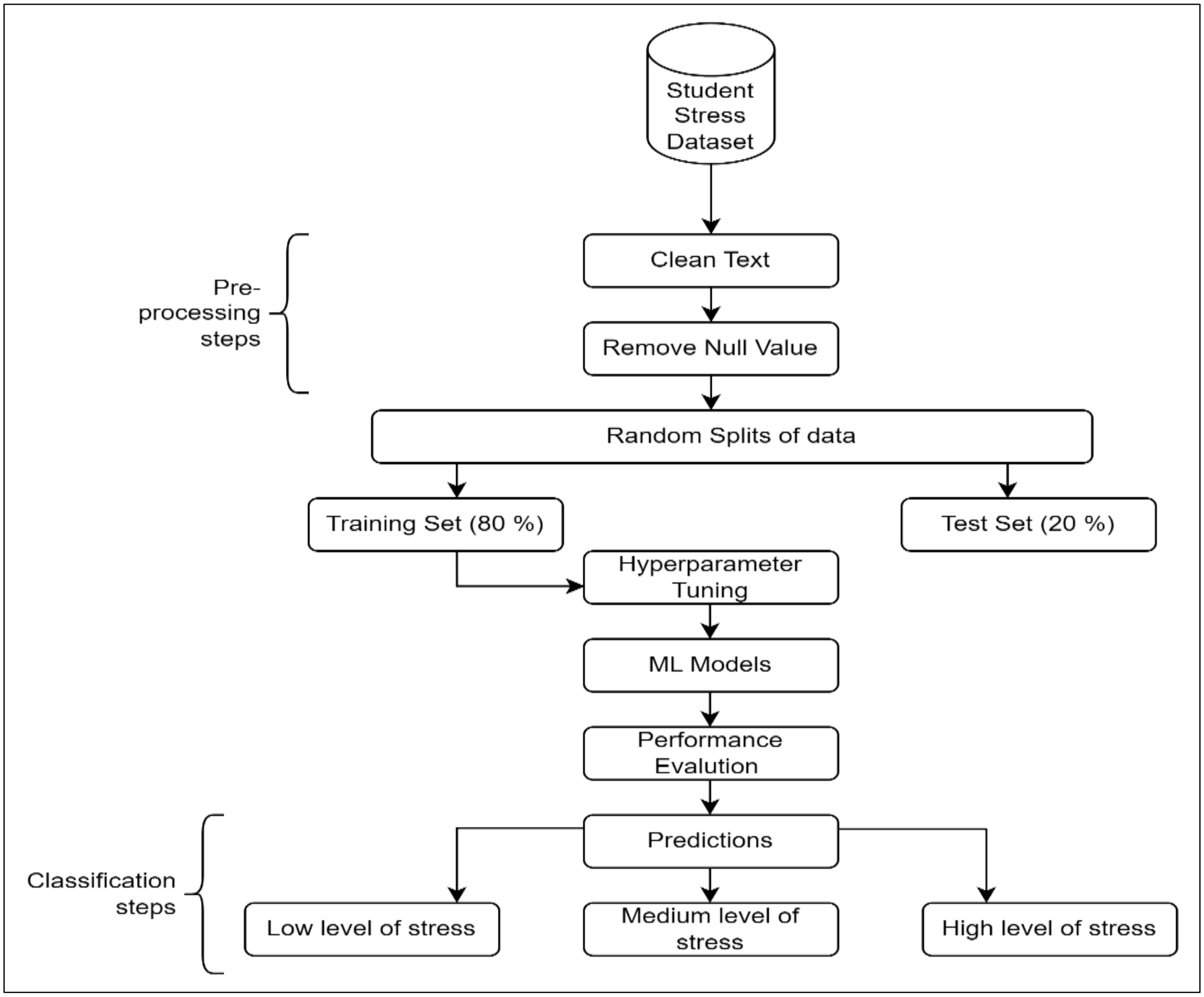

Figure 3 Proposed Methodology

In machine learning, Gradient Boosting (GB) is a boost model for classification and regression activities. Boosting formed ensembles learning where the model has to be trained sequentially, and any new model will try to correct an existing model. It is a combination of weak learners and robust learners. Then, the forecast for a new model is compiled into an ensemble, which will continue until its stop point (Bentéjac et al., 2021). Another ML algorithm used for classification and regression is AdaBoost. AdaBoost is a straightforward algorithm that uses an ensemble technique to improve accuracy. CatBoost (Category Boosting) algorithms improve the original gradient boosting for faster implementation. CatBoost is a supervised ML technique that uses a decision tree for classification and regression. As the name signifies, Cat means category, and Boost means Boosting. It converts categorical variables into numerical variables and is much faster than XGBoost.

LightGBM means Light Gradient Boosting Machine and applies classification techniques for Machine Learning. It is also a decision tree-based learning algorithm and builds a tree based on the histogram method. LightGBM works by combining weak decision trees to make them strong ones. ExtraTree is an ensemble of supervised machine-learning classification algorithms that use decision trees. The Random Forests approach and the ExtraTree algorithm produces different decision trees, but the sampling for every tree is random and non-replaceable. As a result, distinct samples are created in a dataset for every tree. For every tree, a certain quantity of features is also arbitrarily chosen from the entire collection of features (Ampomah et al., 2020).

Extreme gradient boosting (XGBoost) is a supervised decision tree-based machine learning technique that combines decision trees to enhance model accuracy. It is scalable and constructed in parallel (Wade, 2020). The Naive Bayes algorithm assigns equal weight to each feature or characteristic, making it efficient because no property has an effect over another. The algorithm will become efficient as one property has no effect over another. NBC is a simple, proficient, and widely used classification algorithm for text categorization, according to (Chen et al., 2021). Another well-known machine learning method is decision trees, which classify data according to predetermined criteria.

The tree comprises two entities: nodes and leaves denote decisions. A multi-layer perceptron (MLP) is a neural network (NN) with multiple layers. All layers are fully connected to the network. It works in the forward direction only. According to Desai and Shah (2021), MLP uses the backpropagation technique to enhance the model's performance. An artificial neural network is a system composed of several simple processing units operating in parallel. Processing takes place in each node or computing component with a low processing capability; the network's function is defined by its structure and the weight of its connections (Guillod et al., 2020).

3. EXPERIMENTAL EVALUATION AND MODELING OF DL CLASSIFIERS

"Google Colab" is used for implementation, a cloud-based platform provided by Google. Colab contains well-known data science libraries, such as Keras, TensorFlow, PyTorch, and Scikit-learn. The ANN model was trained for ten epochs and obtained an accuracy of 88.63 %. The final ANN model is shown in Figure 4. The ANN architecture follows the pattern of layers, and each layer contains neurons. The neural network's final model's visual representation is displayed in Figure 5. In this section, the following subsections discussed the student stress dataset.

```
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 dense (Dense)               (None, 128)               2688

 dense_1 (Dense)             (None, 64)                8256

 dense_2 (Dense)             (None, 3)                 195

=================================================================
Total params: 11139 (43.51 KB)
Trainable params: 11139 (43.51 KB)
Non-trainable params: 0 (0.00 Byte)
_________________________________________________________________
```

Figure 4 ANN model at optimum performance

In this study, the architecture of MLP contains an input layer, two hidden layers, and an output layer, as shown in Figure 6. For implementation purposes, we used twenty neurons in input and hidden layers.

Datasets

The data set used in this paper is collected from students studying in high schools and colleges of Tribhuvan University, Dharan, Nepal. It has twenty features that can be used to predict and measure their stress levels. This dataset has twenty-one columns and 1100 rows. We have taken "stress_level" as a dependent variable out of twenty-one columns, and other variables are independent. The target variable is a ternary attribute, which provides a stress level diagnosis.

Table 2 shows the student's stress level. The value "two" represents a highly stressful situation, "one" is medium, and "zero" means no stress exists. The data set contains 1,100 rows, of which 880 rows are used to train the model and 220 are used for testing. Independent features, for example, anxiety_level range from 0 to 21, depression (0 to 27), self-esteem (0 to 30), and all others lie between zero to five. In most cases, higher values in anxiety level, depression, and lower values in self-esteem might be associated with higher stress levels. A random sample of the dataset is given in Table 2. The summary statistics of features is given in Table 3.

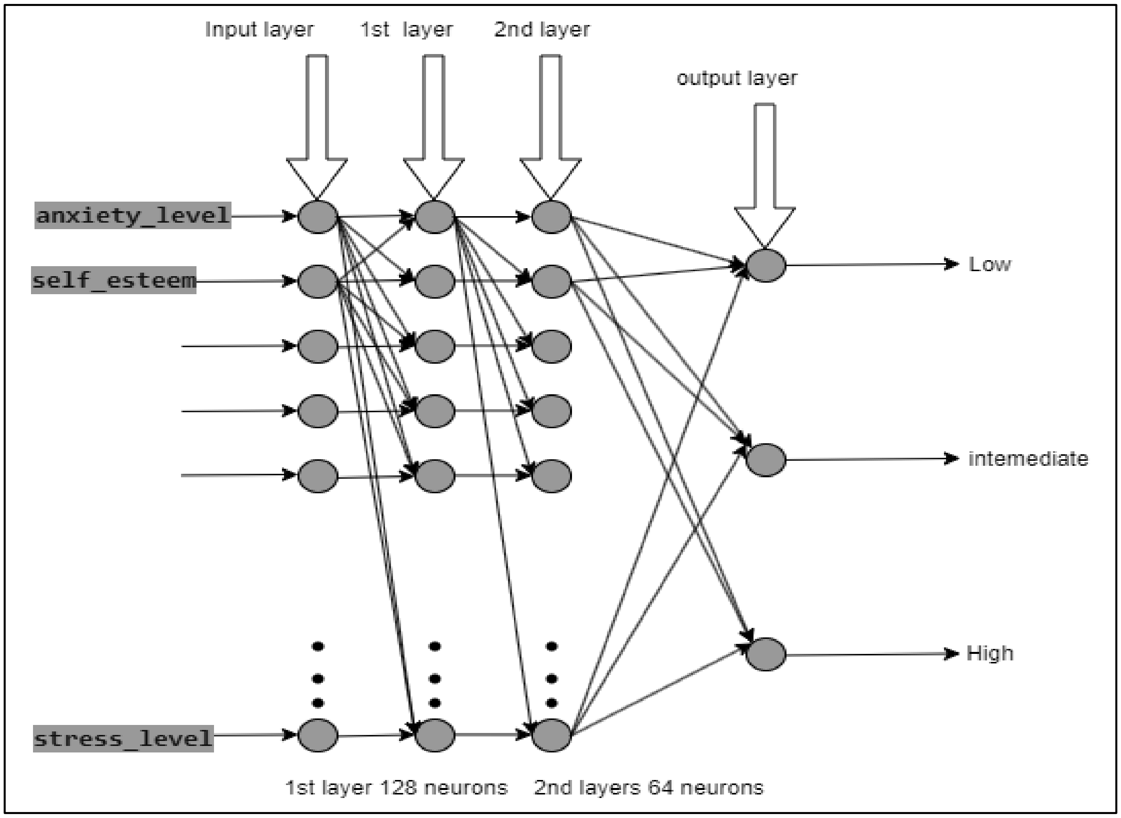

Figure 5 ANN architecture for student stress level

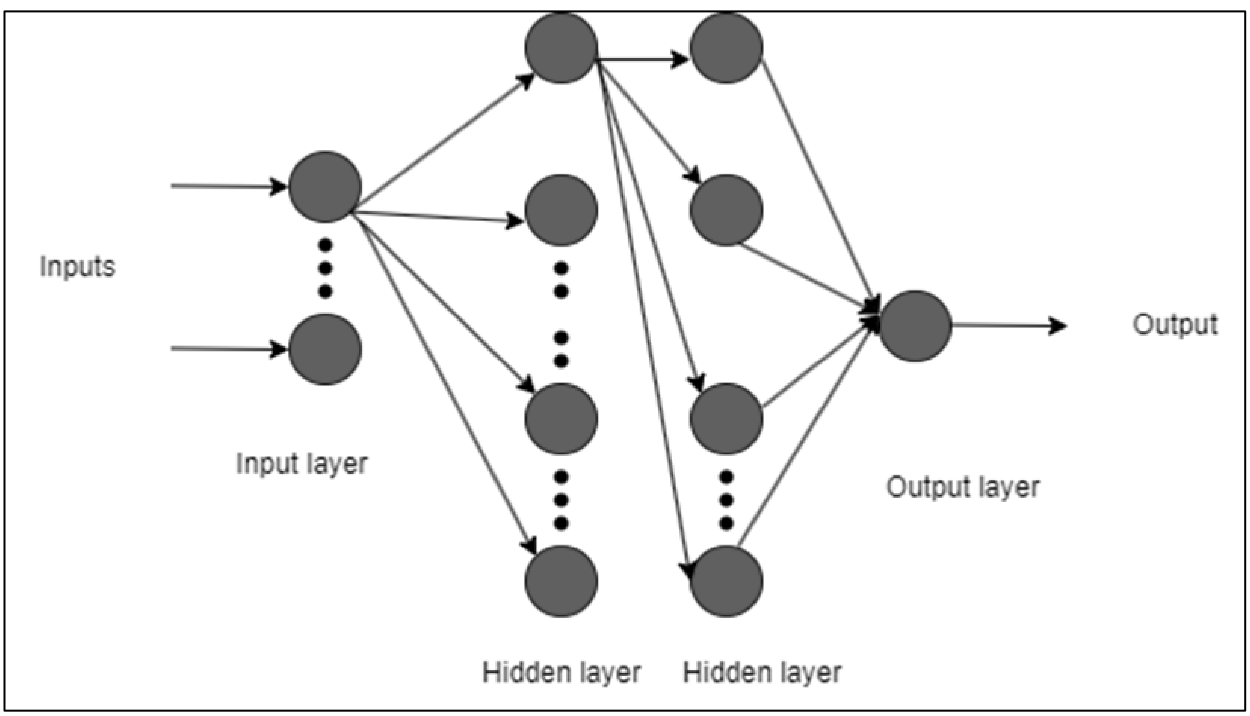

Figure 6 MLP Architecture

Table 2 Sample of Student Stress Dataset

Index of random rows	70	884	1013	808	107
Anxiety_level	5	15	21	18	3
Self_esteem	30	30	8	10	28
Mental_health_history	0	0	1	1	0
Depression	0	7	22	18	1
Headache		3	4	5	1
Blood_pressure	2	3	3	3	2
Sleep_quality	5	4	5	1	4
Breathing_problem	2	0	0	4	1
Noise_level	1	0	5	5	2
Living_conditions	4	0	3	1	4
Basic_needs	4	2	2	2	5
Academic_performance	5	5	5	2	4
Study_load	1	2	1	5	2
Teacher_student_relationship	4	2	4	2	4
Future_career_concerns	1	4	4	4	1
Social_support	3	0	0	1	3
Peer_pressure	1	5	3	4	1
Extracurricular_activities	1	4	4	5	2
Bullying	1	1	2	5	1
Stress_level	0	1	0	2	0

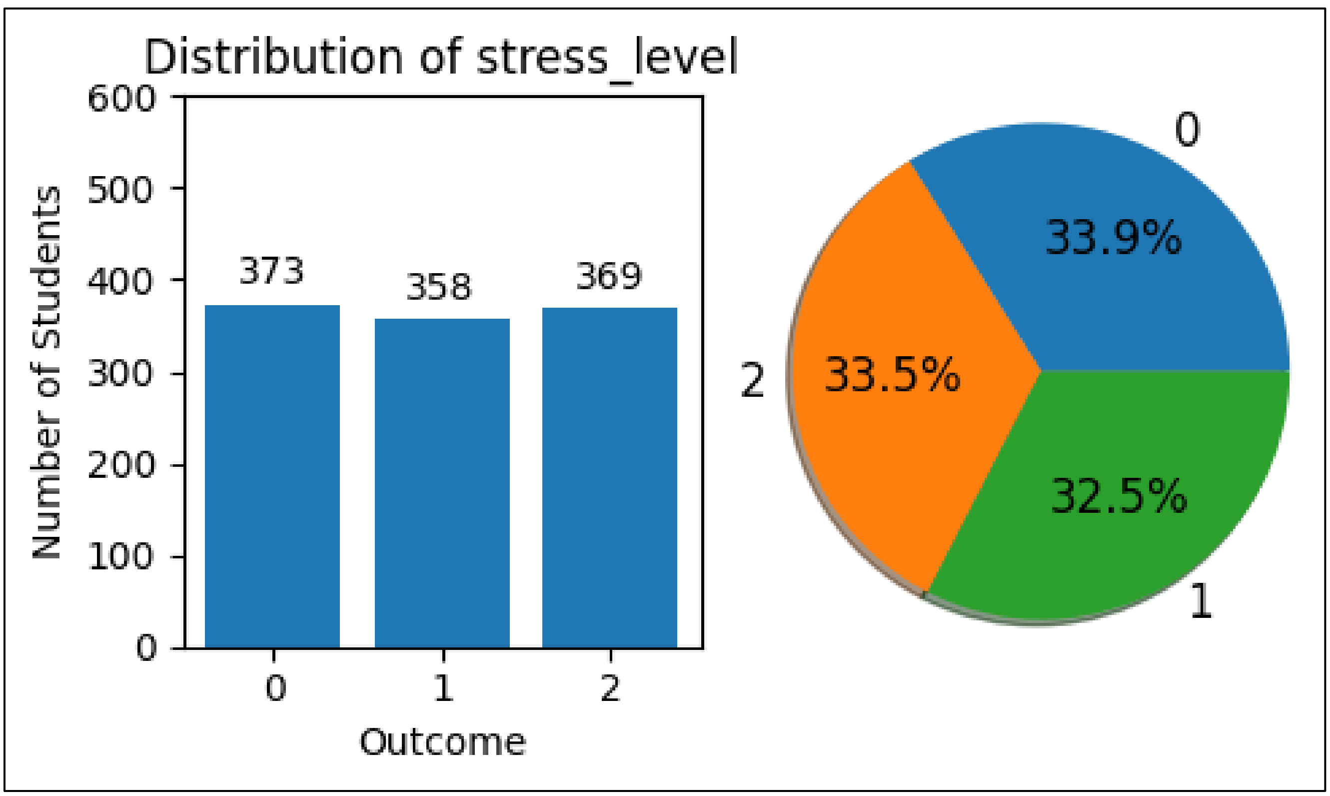

Figure 7 Distribution of student stress level

Table 3 Summary Statistics of features

	Count	Mean	Std	Min	25%	50%	75%	Max
Anxiety_level	1100	11.06364	6.117558	0	6	11	16	21
Self_esteem	1100	17.77727	8.944599	0	11	19	26	30
Mental_health_history	1100	0.492727	0.500175	0	0	0	1	1
Depression	1100	12.55546	7.727008	0	6	12	19	27
Headache	1100	2.508182	1.409356	0	1	3	3	5
Blood_pressure	1100	2.181818	0.833575	1	1	2	3	3
Sleep_quality	1100	2.66	1.548383	0	1	2.5	4	5
Breathing_problem	1100	2.753636	1.400713	0	2	3	4	5
Noise_level	1100	2.649091	1.328127	0	2	3	3	5
Living_conditions	1100	2.518182	1.119208	0	2	2	3	5
Basic_needs	1100	2.772727	1.433761	0	2	3	4	5
Academic_performance	1100	2.772727	1.414594	0	2	2	4	5
Sudy_load	1100	2.621818	1.315781	0	2	2	3	5
Teacher_student_relationship	1100	2.648182	1.384579	0	2	2	4	5
Future_career_concerns	1100	2.649091	1.529375	0	1	2	4	5
Social_support	1100	1.881818	1.047826	0	1	2	3	3
Peer_pressure	1100	2.734545	1.425265	0	2	2	4	5
Extracurricular_activities	1100	2.767273	1.417562	0	2	2.5	4	5
Bullying	1100	2.617273	1.530958	0	1	3	4	5
Stress_level	1100	0.996364	0.821673	0	0	1	2	2

Exploratory Data Analysis

This section explains the stress factors and the correlation between them. "Stress_level" is a multiclass variable, meaning it can be categorized as low, medium, or high. Therefore, it is essential to investigate the impact of psychological, physiological, environmental, social, and academic factors on the "stress level". As discussed, "stress_level" is a multiclass variable, i.e., low, medium, and high, so it is important to find out how psychological, physiological, environmental, social, and academic factors affect it. Thus, correlation matrixes were derived.

The psychological factors correlation matrix has a strong positive correlation between anxiety level and stress level, 0.74. A strong correlation is -0.76. It exists between self-esteem and stress level. A feature such as self-esteem of psychological factors has a negative correlation among all psychological factors, as shown in Figure 8. The physiological factors correlation matrix has a strong positive relation, 0.71, between headache and stress level. A strong negative correlation is -0.64 between sleep quality and stress level.

In Figure 9, sleep quality has negative correlations of -0.75, -0.64, -0.30, and -0.54. The correlation matrix shows that noise level is negatively correlated (−0.57) with basic needs and positively related with stress level, i.e. 0.66. Figure 10 shows the correlation matrix for the environmental factors. With a correlation coefficient of 0.75, as shown in Figure 11, the association between stress levels and bullying is high. Among social factors, social support has a negative correlation with peer pressure, extracurricular activities, and bullying. Figure 12 displays the correlation matrix of academic factors, which shows a positive correlation (0.74) between future_career_concerns and stress level and a strong negative association with academic performance of -0.72.

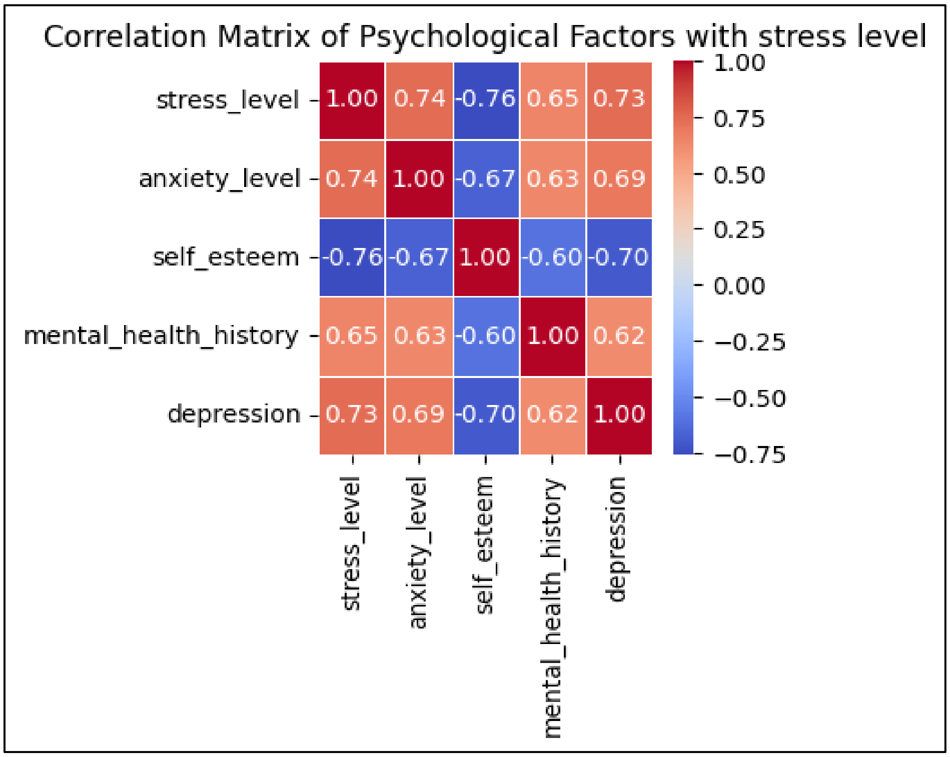

Figure 8 Psychological Factors Correlation Matrix

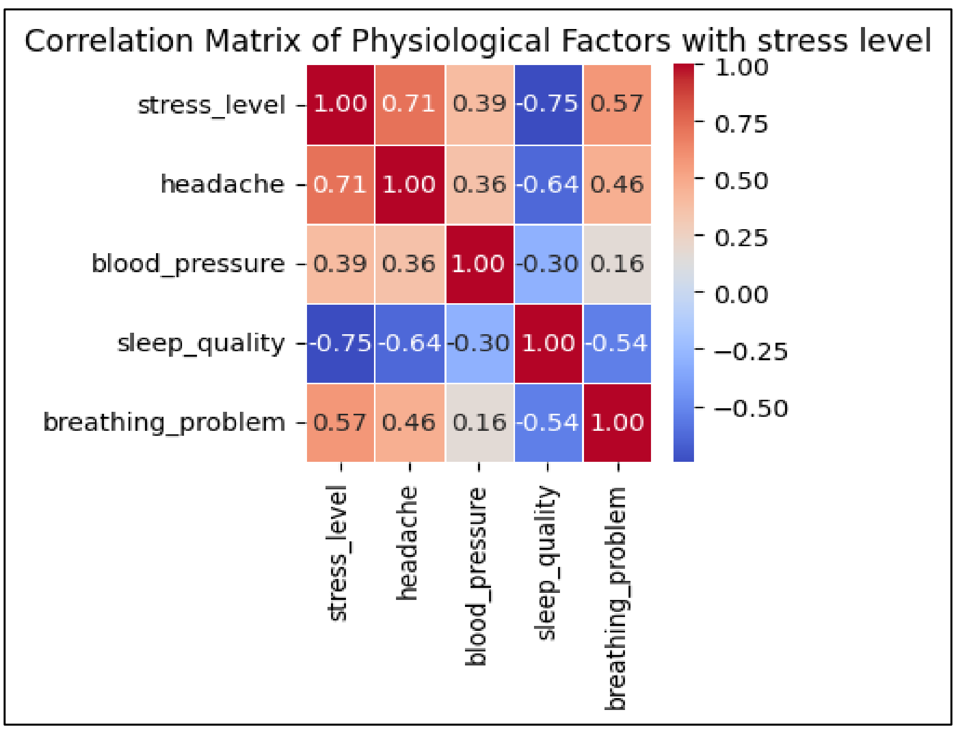

Figure 9 Physiological Factors Correlation Matrix

Figure 13 illustrates the relationships between "stress_level" and all other factors. Figure 14 demonstrates the density distribution of student stress for anxiety level, self-esteem, and depression. The boxplot of anxiety_level for 'no stress (0)' exists between 0 to 14, for 'medium stress (1)', it is between 6 to 17, and for 'high stress (2), it is 9 to 21, as shown in Figure 14. This category also has a few outliers below the lower and upper margins. Most students' self-esteem lies between 0 and 25 for high-stress levels, as shown in Figure 14.

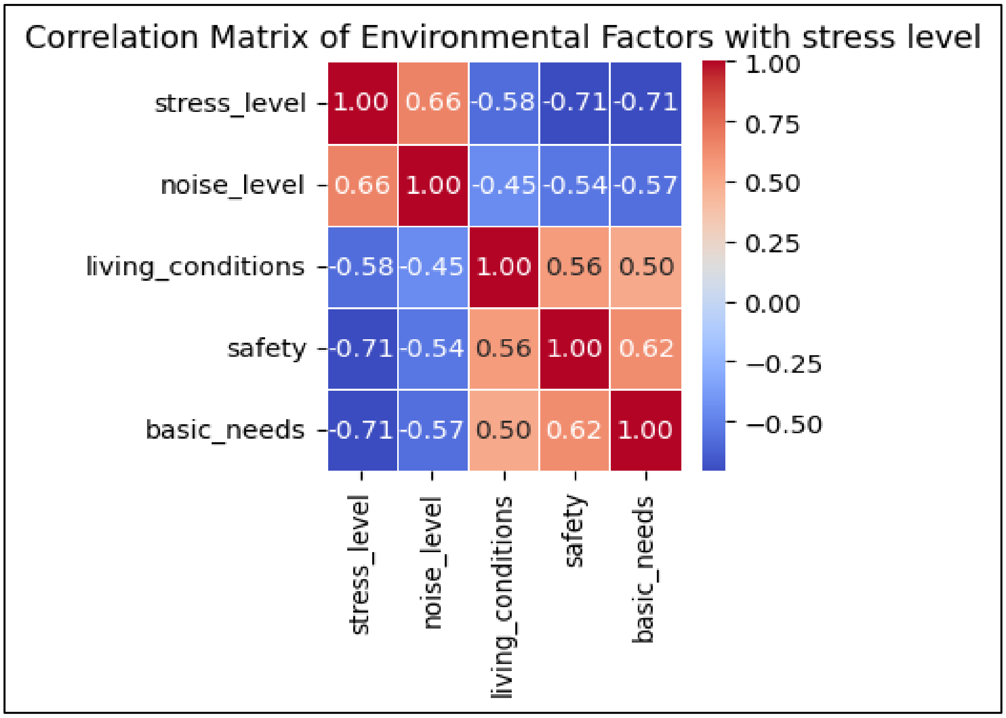

Figure 10 Environmental Factors Correlation Matrix

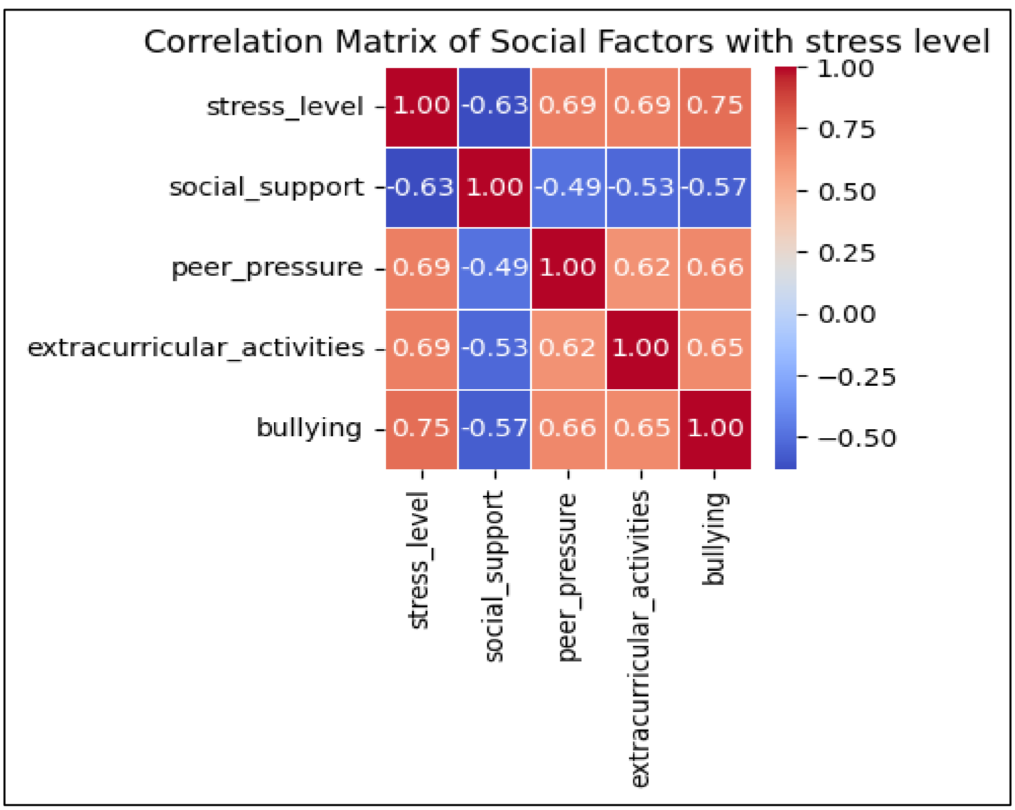

Figure 11 Social Factors Correlation Matrix

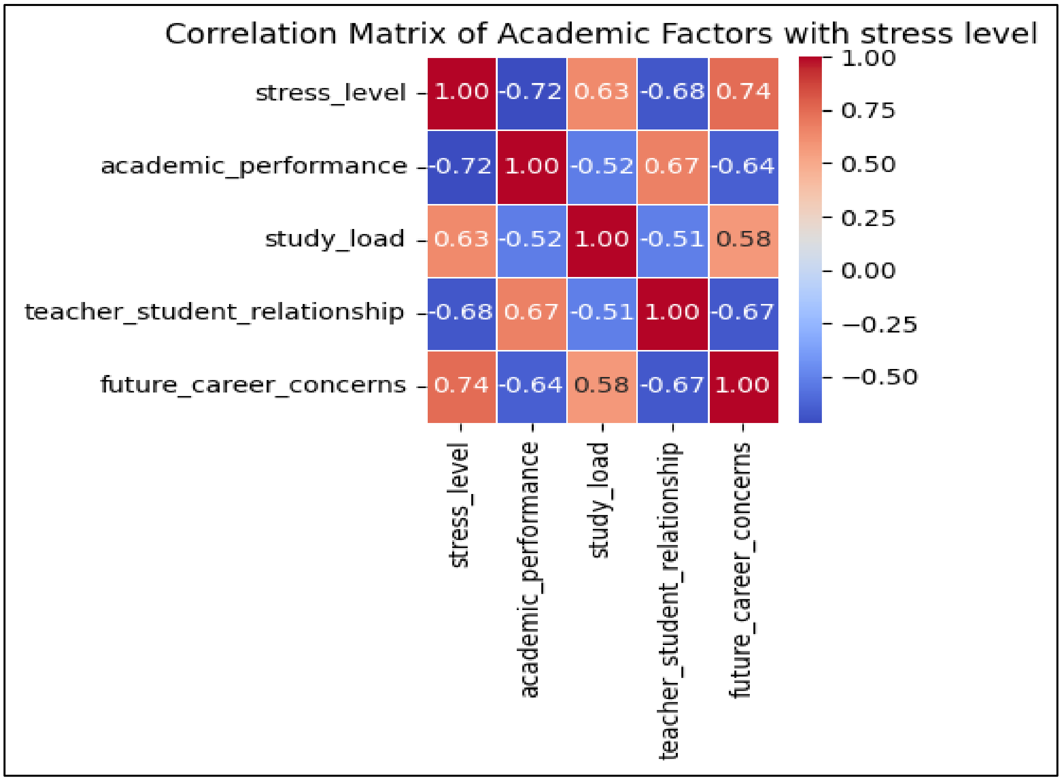

Figure 12 Academic Factors Correlation Matrix

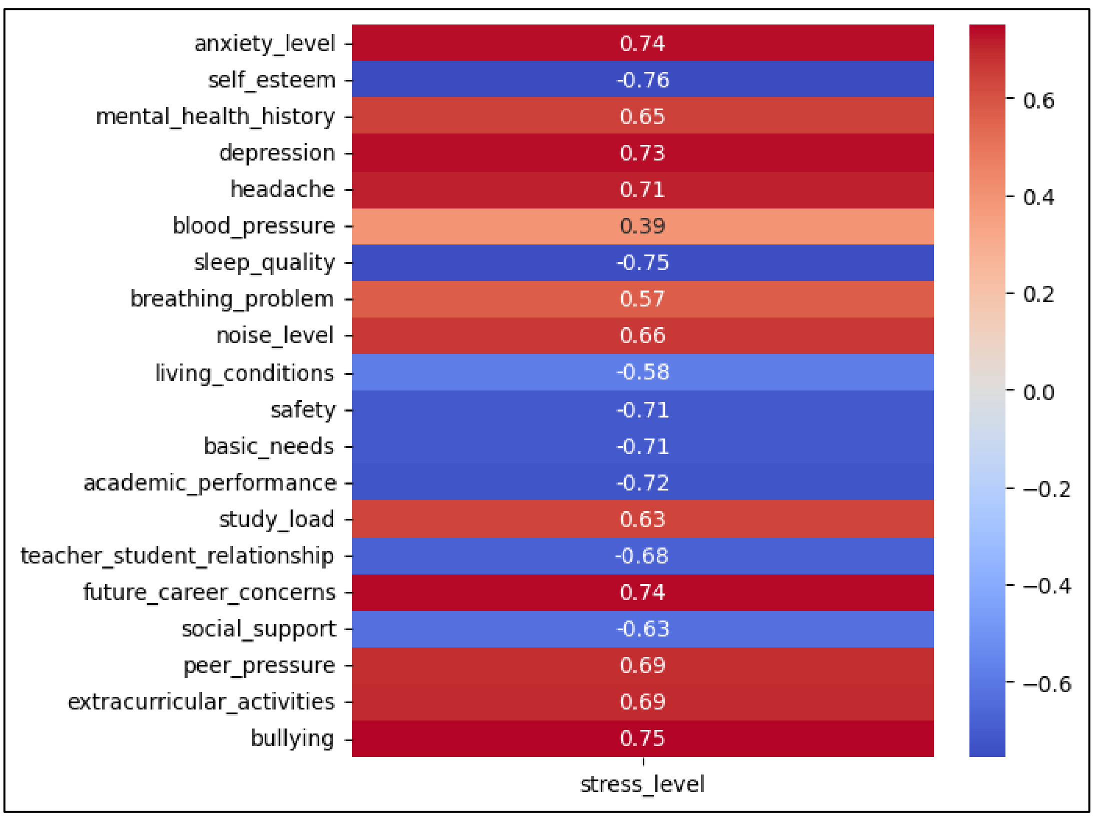

Figure 13 Correlation of stress level with all factors

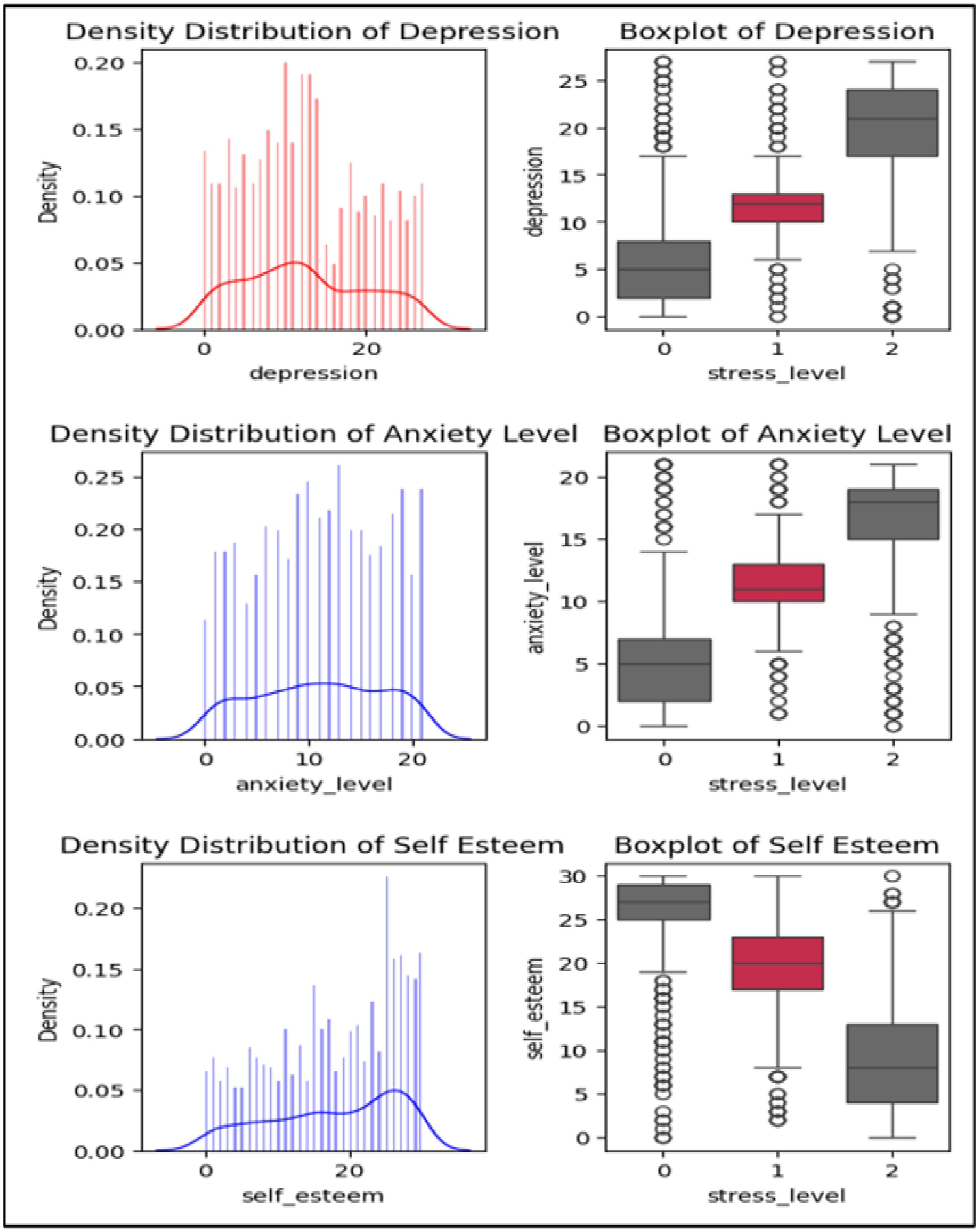

Figure 14 Density and Boxplot of Factors anxiety_level, self-esteem, and depression

Performance Evaluation

The performance of the ML model is evaluated through recall, precision, accuracy, and F1-score. These are based on different components such as True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN). FP is a Type I Error, and FN is a Type II Error. TP: When both (i.e., actual and predicted values) are one. TN: When both (i.e., actual and predicted) are zero. FP: When actual is zero and predicted is one. FN: When actual and predicted values are one and zero, respectively.

Definitions and formula

Accuracy: Accuracy is the sum of true positive and true negative divided by the total number of classifieds, which is:

Accuracy = (TP + TN) ÷ (TP + TN + FP + FN)
Precision: When the true positive value of the confusion matrix is divided by a combination of TP and FP, then:
Precision (Specificity) = TP ÷ (TP + FP)
Sensitivity is also known as recall, and it can be obtained by dividing TP by the total number of TP and FN.
Sensitivity = TP ÷ (TP + FN)
F1-score is the average of recall and precision, which is more critical than accuracy.
F1-score = 2 ∗ ((precision ∗ recall) ÷ (precision + recall))
The shapes of the training dataset have 880 rows and 20 columns, while test and validation contain 220 rows and 20 columns, as shown in Table 4. In this paper, ML techniques like GB, Adaptive Boosting (AdaBoost), LightGBM (LGBM), RF, CatBoost, SVM, ExtraTree (ET), XGBoost, Linear Regression, LR, KNN, NB, and DT are used for classification and regression. DL models like ANN and MLP are also used in this paper to predict stress levels.

Table 4 Shape of Student Stress Dataset

Dataset	**Shape**
Training Set	(880,20)
Validation Set	(220,20)
Test Set	(220,20)

Employing the right technique for machine-learning and deep-learning is essential in developing a highly precise and reliable classifier. The results of hyperparameter optimization for the ML models ExtraTree (ET), Random Forest (RF), AdaBoost, and Gradient Boosting (GB) over ten folds are presented in Table 5. Varied parameters were utilized for hyperparameter tuning. The Gradient Boosting model achieved an accuracy of 0.8897 with the optimal parameters (learning_rate: 1, max_depth: 7, and n_estimators: 500). For the AdaBoost model, the best parameters were found to be learning_rate: 0.25 and n_estimators: 500, resulting in an accuracy of 0.8852.

Similarly, the Random Forest model attained its highest accuracy of 0.8988 with the following parameters: max_depth: None, max_features: 1.0, max_samples: 0.75, and n_estimators: 20. Lastly, the ExtraTree model demonstrated an accuracy of 0.8863 with the best parameters: Criterion: Gini, min_samples_leaf: 4, max_depth: 20, min_samples_split: 4, and n_estimators: 100.

Table 5 Hyperparameter optimization result for ML.

Model	**Parameters**	**Best Parameters**	**Accuracy**
Extra Tree	{'n_estimators': [100, 105, Ellipsis, 500], 'max_depth': [5, 10, 15, 20], 'criterion': ['gini','entropy'], 'min_samples_split': [2, 4, 6], 'min_samples_leaf': [4, 5, 6]}	{'criterion': 'gini', 'max_depth': 20, 'min_samples_leaf': 4, 'min_samples_split': 4, 'n_estimators': 100}	0.8863
Random Forest	{'n_estimators': [20, 100, 130], 'max_features': [0.2, 0.6, 1.0], 'max_depth': [10, 12,	{'max_depth': None, 'max_features': 1.0, 'max_samples': 0.75, 'n_estimators': 20}	0.8988

	None], 'max_samples': [0.5, 0.75, 1.0]}		
AdaBoost	{'n_estimators': [100, 105, Ellipsis, 500], 'learning_rate': [0.25, 0.5, 0.75, 0.9]}	{'learning_rate': 0.25, 'n_estimators': 500}	0.8852
Gradient Boosting	{'n_estimators': [100, 105, 500], 'max_depth': [-1, 3, 7, 14, 21], 'learning_rate': [0.01, 0.1, 0.5, 1]}	{'learning_rate': 1, 'max_depth': 7, 'n_estimators': 500}	0.8897

4. RESULTS AND DISCUSSION

This study aimed to analyze the levels of PPAES components of student stress. In this study, we have developed machine learning and deep learning models to predict different stress states with the help of PPAES factors. Various ML techniques were trained, and confusion matrix and classification reports were obtained. Figures 15 and 16 show the multiclass RF and ANN model performance using the ROC curve. Under a high-stress level, the ANN ROC value is 0.99, which shows better accuracy. Table 6 displays the NB algorithm's recall, precision, accuracy, and support during low, intermediate, and high-stress phases.

Table 6 Classification Report of NB

	Precision	Recall	f1-score	Support
0	0.96	0.89	0.93	76
1	1.00	0.84	0.91	73
2	0.78	0.97	0.87	71
Accuracy	-	-	0.90	220
Macro avg	0.91	0.90	0.90	220
Weighted avg	0.92	0.90	0.90	220

The model's performance is evaluated by comparing the F1 score and precision. The Naïve Bayes model achieved 96% precision in the 'no stress' scenario and 78% accuracy in the 'high stress' scenario. Overall, the Naïve Bayes model has an accuracy of 90%. The x-axis represents the model-predicted label, and the y-axis represents the actual label. To compare different models, we assess how well they predict true positives (TP) and true negatives (TN). We select the model that performs better in predicting TP and TN as our base model. For instance, the Naïve Bayes model has TP=68, TN=61,11,0,69, FN=0, 8 and FP=1, 2, as depicted in Figure 19B.

The ROC curve illustrates the model's performance at various thresholds. The y-axis represents sensitivity, and the x-axis represents the false positive rate. The AUC for the RF model with the highest stress level falls under Class 2, as shown in Figure 15. The loss and accuracy curve for the ANN model is depicted in Figure 17, covering ten epochs. Additionally, Figure 18 displays the accuracy curve for the MLP, showing both training and testing accuracy.

In the first row, there were 76 instances of class 0, of which 68 were correctly identified as class 0. In the second row, there were 73 instances of class 1, with the classifier correctly identifying 61. In the confusion matrix's third row, 71 instances, out of which 69 belonged to class 2, as shown in Figure 19B. By utilizing the above confusion matrices, we can calculate the precision, recall, and f1-score for high (2), medium (1), and low (0) stressful situations, as shown in Table 7.

The state-of-the-art models demonstrated by Rescio et al., (2024) describe the stress of twenty workers through 1D-CNN, LSTM, and GRU models of deep learning and obtained an accuracy of 95.38% (1D-CNN). The dataset used in this paper is not publicly available. A stress detection framework based on ML and IoT was proposed by (Bansal and Vyas, 2024). Stress detection using the proposed MLIoT-ESD technique is more time-consuming as compared to traditional approaches. The above-mentioned papers do not adopt the ML, DL models, and dataset, which are implemented by the authors of this paper. The comparative results of various machine learning techniques are presented in Table 7. The table illustrates the performance of machine learning techniques, including SVM, XGBoost, AdaBoost, Random Forest, LightGBM, Gradient Boosting, decision tree, CatBoost, ExtraTree, KNN, Naïve Bayes, and logistic regression, based on calculated train accuracy, test accuracy, recall, precision, and F1-score of each classifier.

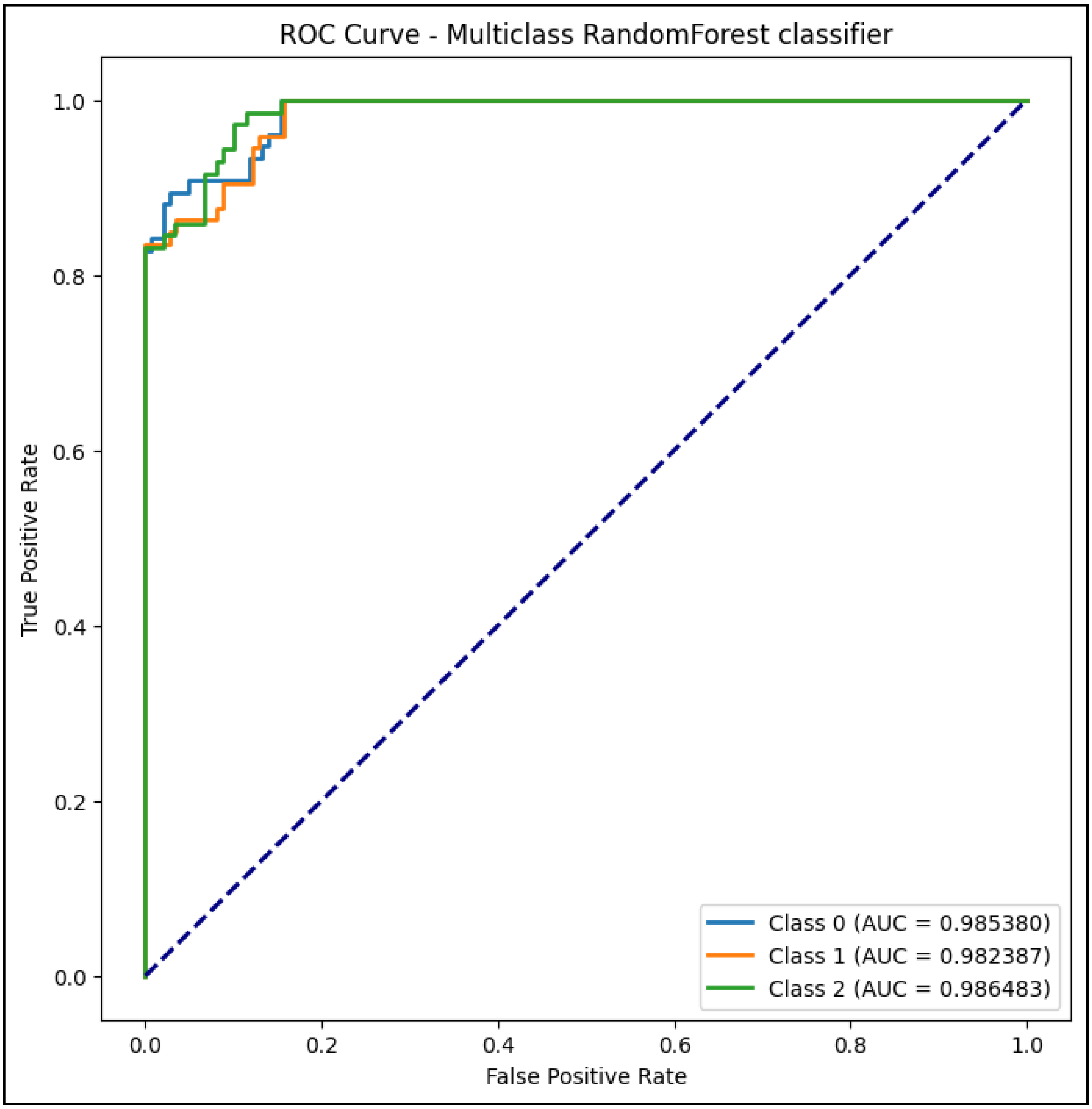

Figure 15 ROC curve of RF model

These findings demonstrate the superiority of the suggested techniques across a wide range of established classifiers, underscoring their potential for practical applications. Additionally, the results of deep learning techniques, specifically MLP and ANN, are displayed in Table 8, showcasing test accuracy (90%), recall (93%), precision (92%), and F1-score (92%). When comparing deep learning and machine learning classifiers, MLP demonstrates better precision, recall, and F1-score accuracy of 90%. Naïve Bayes produced the same test accuracy (0.9000) as MLP, but its precision (0.7841), recall (0.9718), and F1-score (0.8679) are not as strong when compared to the deep learning classifier. Among all the machine learning models, Naïve Bayes has the highest performance.

Random Forest (RF) and Naïve Bayes exhibit similar recall and test accuracy, but their F1-score and precision differ. Similarly, AdaBoost and CatBoost display equivalent precision, recall, and F1 scores. The top-performing classifier models are MLP, Naïve Bayes, Random Forest, ANN, and LightGBM. Conversely, lower-performing classifiers such as CatBoost, Decision Tree, KNN, and SVM demonstrate lower accuracy, recall, and F1 scores. Decision Tree and KNN have identical test accuracy, with a slight difference in the F1-score. Overall, this indicates that the choice of a classifier can significantly impact the performance of prediction models. In addition, Figure 19 presents the confusion matrix of machine learning models, accurately predicting students' stress levels. This study was conducted to evaluate the impact of stress on students' performance across various domains.

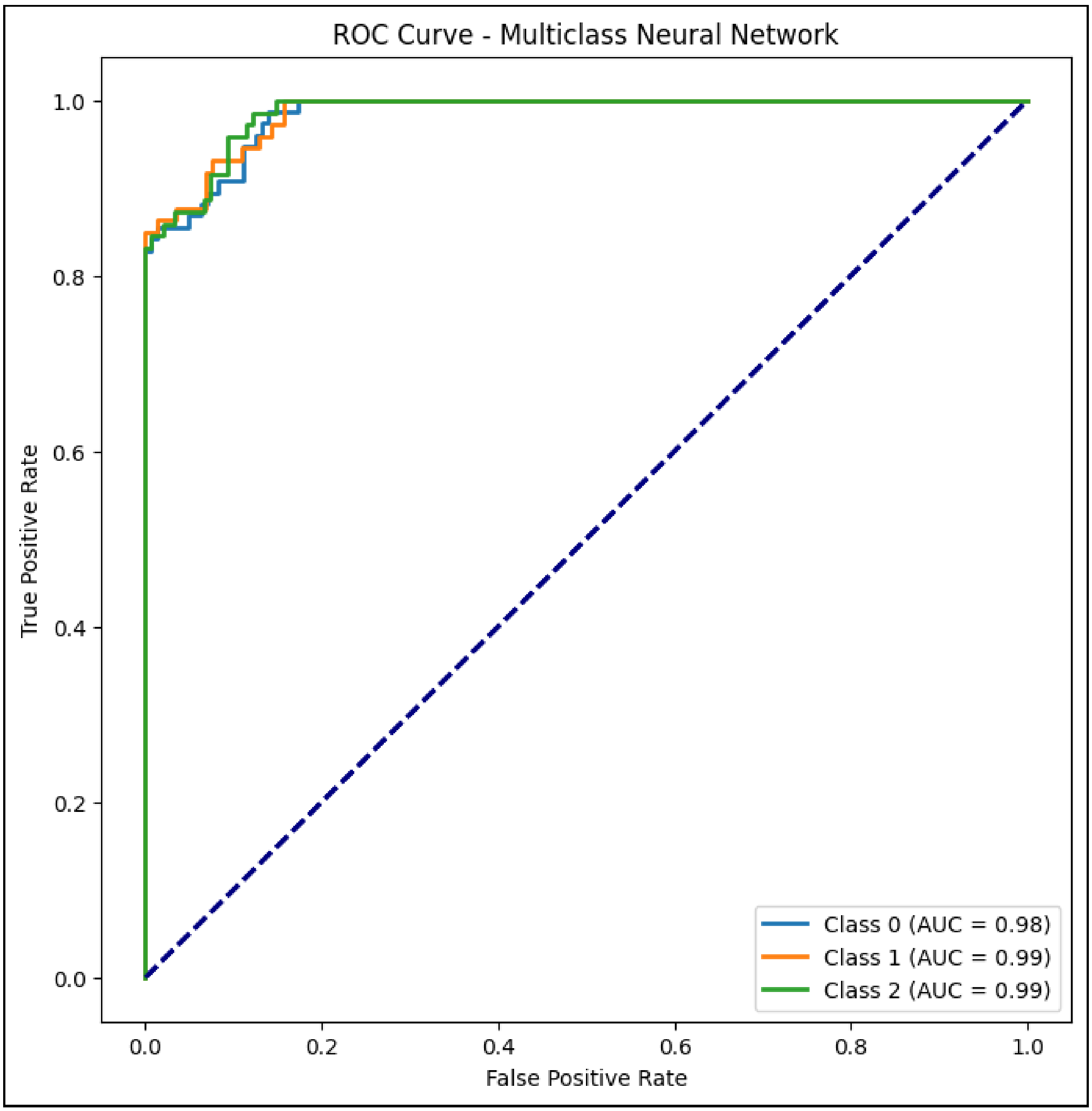

Figure 16 ROC curve of ANN model

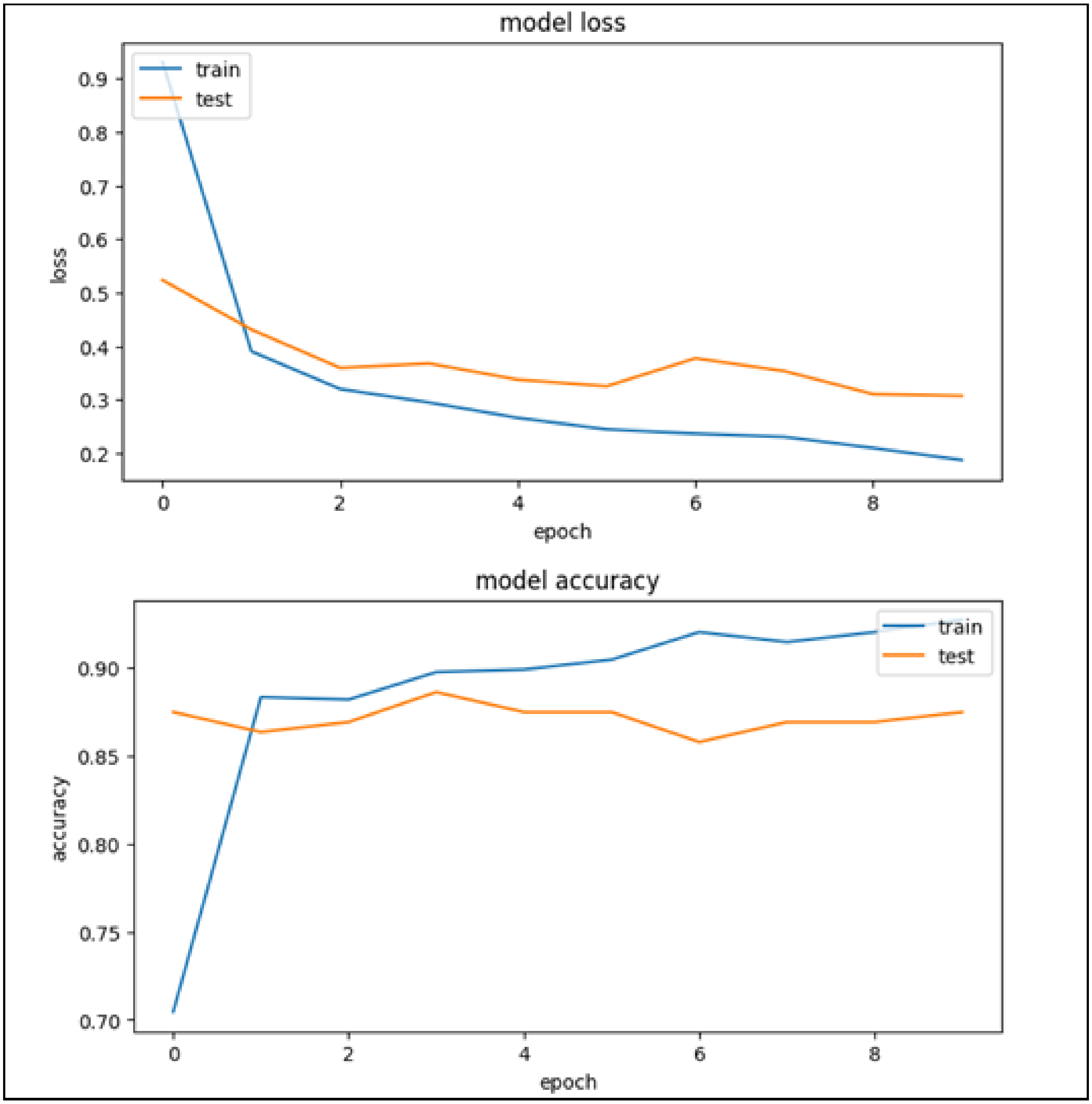

Figure 17 ANN loss and accuracy curve

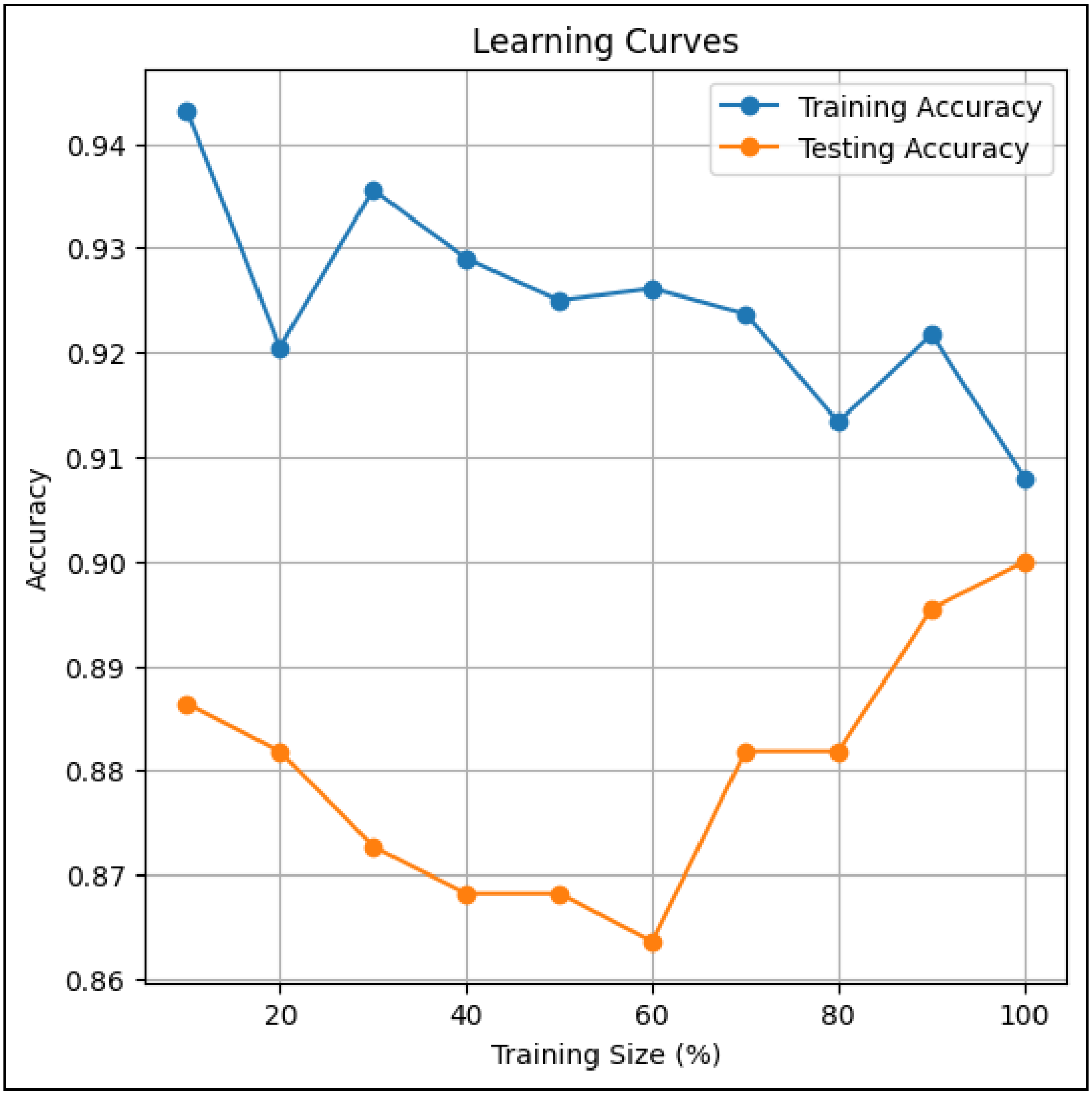

Figure 18 MLP accuracy curve

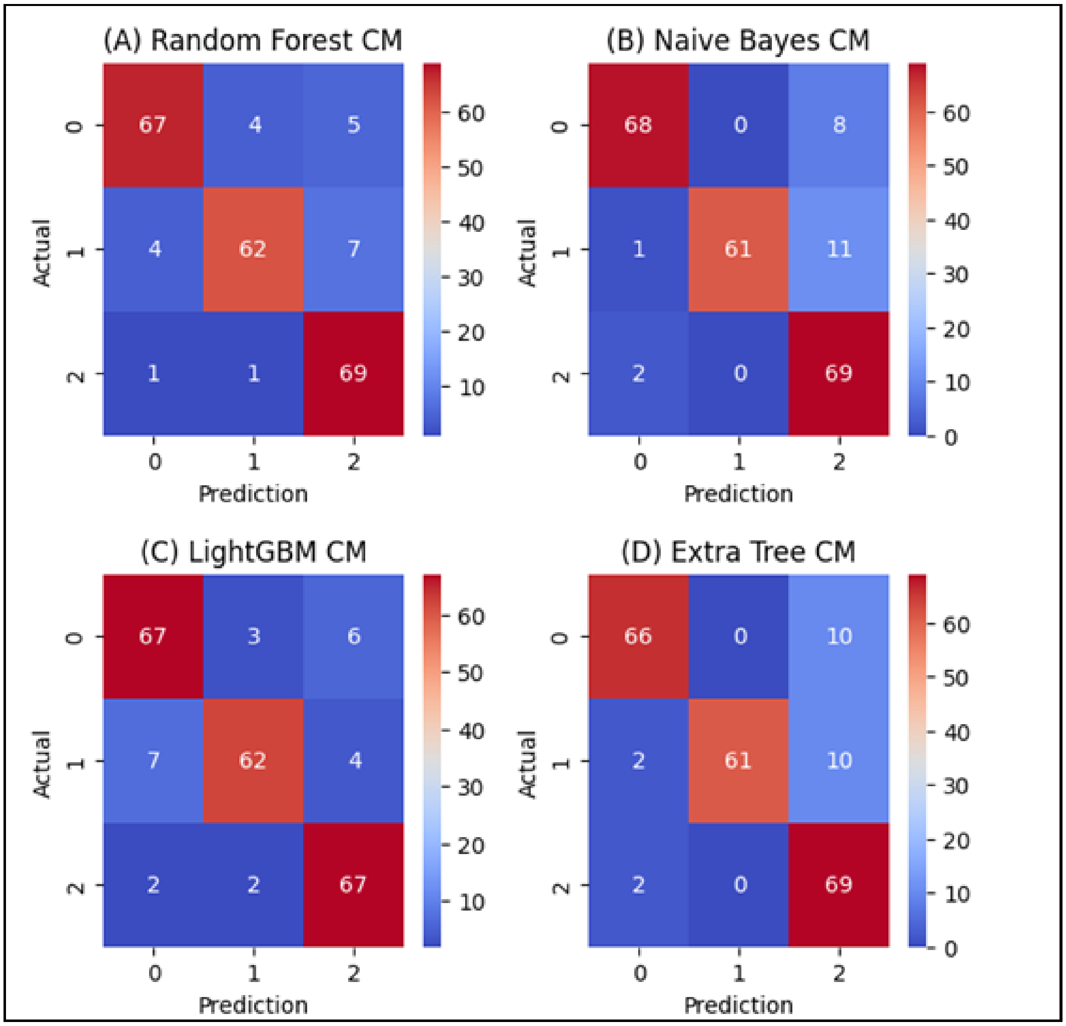

Figure 19 Machine learning model's confusion matrix

Table 7 Results of machine learning models

Classifier	Train Accuracy	Test Accuracy	Precision	Recall	f1-score
Naïve Bayes	0.8784	0.9000	0.7841	0.9718	0.8679
RF	0.9170	0.9000	0.8519	0.9718	0.9079
LightGBM	0.9080	0.8909	0.8701	0.9437	0.9054
ExtraTree	0.8920	0.8909	0.7753	0.9718	0.8625
XGBoost	0.9023	0.8864	0.8590	0.9437	0.8993
GB	0.9136	0.8818	0.8800	0.9296	0.9041
Logistic Regression	0.9068	0.8818	0.9143	0.9014	0.9078
AdaBoost	0.9307	0.8773	0.8873	0.8873	0.8873
CatBoost	0.9170	0.8682	0.8873	0.8873	0.8873
Decision Tree	0.9420	0.8591	0.9062	0.8169	0.8593
KNN	0.9068	0.8591	0.8824	0.8451	0.8633
SVM	0.9795	0.8545	0.8806	0.8310	0.8551

Table 8 Results of deep learning Models

Classifier	Test Accuracy	Precision	Recall	f1-score
MLP	0.9000	0.9200	0.9300	0.9200
ANN	0.8954	0.9000	0.9000	0.9000

5. CONCLUSION

The following text discusses the final findings of the stress factor identification models. The study uses a combination of machine learning (ML) and deep learning (DL) to analyze the dataset. These models can identify various factors responsible for causing stress. Students facing stress can be identified using models such as SVM, Random Forest, GB, AdaBoost, CatBoost, LightGBM, ExtraTree, XGBoost, LR, KNN, NB, DT, Multi-Layer Perceptron (MLP), and Artificial Neural Network (ANN). The accuracy of the Naïve Bayes model is 90%, while SVM has the lowest test accuracy level at 85.45%. The study reveals that the academic period is a critical factor of stress for students.

Physiological and psychological factors show moderate stress levels among college and school-going students. Social and environmental factors are reported to cause the lowest stress. Based on the study findings, a stress diagnosis system can be developed for students, ultimately enhancing their performance. This research can also be extended in the future by incorporating IoT. Wearable devices for stress detection among students may be developed and embedded with a mobile application to process real-time information.

Acknowledgement
We thank the participants who contributed to conducting this research analysis. We are also thankful to our institutes, which provide various facilities, such as labs and study material.

Author Contributions
Individual contributions to prepare the following research articles is as follows: Conceptualization: Dr Suraj Arya; Methodology Dr Suraj Arya; Data source / Data set: Dr Suraj Arya; Problem Statement: Dr Suraj Arya; Software Dr Suraj Arya & Anju; validation Dr Suraj Arya and Anju; Formal analysis, Dr Suraj Arya and Anju; investigation Dr Suraj Arya and Anju; Resources, Dr Suraj Arya; Writing Dr Suraj Arya and Anju; preparation, Dr Suraj Arya; writing—review and editing, Dr Suraj Arya & Nor Azuana Ramli.; Visualization, Anju & Nor Azuana Ramli.; Supervision, Dr Suraj Arya; Project administration: Dr Suraj Arya. All authors have read and agreed to the published version of the manuscript.

Ethical issues
Not applicable.

Abbreviations
AdaBoost– Adaptive Boosting
ANN- Artificial Neural Network
CatBoost– Category Boosting
DL- Deep Learning
DT– Decision Tree
ET– Extra Tree
FN– False Negative
FP– False Positive
GB– Gradient Boosting
KNN– K-Nearest Neighbors
LightGBM– Light Gradient Boosting

LR– Logistic Regression
ML- Machine Learning
MLP– Multi-Layer Perceptron
NB– Naïve Bayes
PPAES - Physiological, Psychological, Academic, Environmental, and Social
RF- Random Forest
SVC– Support Vector Classifier
SVM- Support Vector Machine
TN- True Negative
TP– True Positive
XGBoost– Extreme Gradient Boosting

Informed consent

Not applicable.

Funding

This study has not received any external funding.

Conflict of Interest

The author declares that there are no conflicts of interests.

Data and materials availability

All data associated with this study are present in the paper.

REFERENCES

1. Al-Atawi AA, Alyahyan S, Alatawi MN, Sadad T, Manzoor T, Farooq-i-Azam M, Khan ZH. Stress Monitoring Using Machine Learning, IoT and Wearable Sensors. Sensors (Basel) 2023; 23(21):8875. doi: 10.3390/s23218875
2. American Psychological Association. Stress. American Psychological Association, 2023.
3. Ampomah EK, Qin Z, Nyame G. Evaluation of Tree-Based Ensemble Machine Learning Models in Predicting Stock Price Direction of Movement. Information 2020; 11(6):332. doi: 10.3390/info11060332
4. Bansal M, Vyas V. Evolutionary Stress Detection Framework through Machine Learning and IoT (MLIoT-ESD). Bentham Science Publishers Ltd. in Recent Pat Eng 2024; 18(8):1-12. doi: 10.2174/0118722121267661231013062252
5. Bentéjac C, Csörgő A, Martínez-Muñoz G. A comparative analysis of gradient boosting algorithms. Artif Intell Rev 2021; 54(3):1937-1967. doi: 10.1007/s10462-020-09896-5
6. Chen H, Hu S, Hua R, Zhao X. Improved naive Bayes classification algorithm for traffic risk management. EURASIP J Advances in Signal Process 2021; 2021(1). doi: 10.1186/s13634-021-00742-6
7. Desai M, Shah M. An anatomization on Breast Cancer Detection and Diagnosis employing Multi-layer Perceptron Neural Network (MLP) and Convolutional Neural Network (CNN). Clinical eHealth 2021; 4(6). doi: 10.1016/j.ceh.2020.11.002
8. Ding C, Zhang Y, Ding T. A systematic hybrid machine learning approach for stress prediction. PeerJ Comput Sci 2023; 9:e1154. doi: 10.7717/peerj-cs.1154
9. Edwards A, Edwards C, Abendschein B, Espinosa J, Scherger J, Vander-Meer P. Using robot animal companions in the academic library to mitigate student stress. Library Hi Tech 2020; ahead-of-print No. ahead-of-print. doi: 10.1108/LHT-07-2020-0148
10. Guillod T, Papamanolis P, Kolar JW. Artificial Neural Network (ANN) Based Fast and Accurate Inductor Modeling and Design. IEEE Open J Power Electron 2020; 1:284–299.
11. Liao Z, Fan X, Ma W, Shen Y. An Examination of Mental Stress in College Students: Utilizing Intelligent Perception Data and the Mental Stress Scale. Math 2024; 12(10):1501. doi: 10.3390/math12101501

12. Liu Y, Wang Y, Zhang J. New Machine Learning Algorithm: Random Forest. In: Liu B, Ma M, Chang J, Eds., Information Computing and Applications, Lecture Notes in Computer Science, 2012; 7473:246–52. doi: 10.1007/978-3-642-34062-8_32
13. MedlinePlus. Stress and your health. Medlineplus 2022.
14. Miao M, Zheng L, Xu B, Yang Z, Hu W. A multiple frequency bands parallel spatial–temporal 3D deep residual learning framework for EEG-based emotion recognition. Biomed Signal Process Control 2023; 79(2):104141. doi: 10.1016/j.bspc.2022.104141
15. Mind. What Is Stress? Mind 2022.
16. Mittal S, Mahendra S, Sanap V, Churi P. How can machine learning be used in stress management: A systematic literature review of applications in workplaces and education. Int J Inf Manag Data Insights 2022; 2(2):100110. doi: 10.1016/j.jjimei.2022.100110
17. Mohamed AA, 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. doi: 10.1016/j.compbiomed.2024.108918
18. Nijhawan T, Attigeri G, Ananthakrishna T. Stress detection using natural language processing and machine learning over social interactions. J Big Data 2022; 9(1):1-24. doi: 10.1186/s40537-022-00575-6
19. Onim SH, Thapliyal H, Rhodus EK. Utilizing Machine Learning for Context-Aware Digital Biomarker of Stress in Older Adults. Information 2024; 15(5):274. doi: 10.3390/Info15050274
20. Ratul IJ, Nishat MM, Faisal F, Sultana S, Ahmed A, Al-Mamun MA. Analyzing Perceived Psychological and Social Stress of University Students: A Machine Learning Approach. Heliyon 2023; 9(6):e17307. doi: 10.1016/j.heliyon.2023.e17307
21. Rescio G, Manni A, Ciccarelli M, Papetti A, Caroppo A, Leone A. A Deep Learning-Based Platform for Workers' Stress Detection Using Minimally Intrusive Multisensory Devices. Sensors (Basel) 2024; 24(3):947. doi: 10.3390/s24030947
22. Richer R, Koch V, Abel L, Hauck F, Kurz M, Ringgold V, Müller V, Küderle A, Schindler-Gmelch L, Eskofier BM, Rohleder N. Machine learning-based detection of acute psychosocial stress from body posture and movements. Sci Rep 2024; 14(1):8251. doi: 10.1038/s41598-024-59043-1
23. Tang Y, Wang Y, Zhang X, Wang Z. STILN: A novel spatial-temporal information learning network for EEG-based emotion recognition. Biomed Signal Process Control 2023; 85: 104999–9. doi: 10.1016/j.bspc.2023.104999
24. Vallone F, Cattaneo- Della-Volta MF, Mayor-Silva LI, Monroy AM, Galletta M, Curcio F, Zurlo MC. The COVID-19 Student Stress Questionnaire: Validation in Spanish university students from health sciences. Health Psychol Open 2022; 9(2):205510292211352. doi: 10.1177/20551029221135293
25. Wade C. Hands-On Gradient Boosting with XGBoost and scikit-learn: Perform accessible machine learning and extreme gradient boosting with Python. Packt Publishing; 2020.
26. Zhang C, Shao X, Li D. Knowledge-based Support Vector Classification Based on C-SVC. Procedia Comput Sci 2013; 17: 1083–90. doi: 10.1016/j.procs.2013.05.137