ISSN: 2584-0495 Vol. 4, Issue 1, pp. 1838-1843

International Journal of Microsystems and IoT

ISSN: (Online) Journal homepage: https://www.ijmit.org

Prediction of Mental Stress, Depression & Suicidal Symptoms in Students using Machine Learning Techniques

Sanjiv Sinha, Indrajeet Mukherjee

Cite as: Sinha, S., & Mukherjee, I. (2026). Prediction of Mental Stress, Depression & Suicidal Symptoms in Students using Machine Learning Techniques. International Journal of Microsystems and IoT, 4(1), 1838–1843. https://doi.org/10.5281/zenodo.18409522

https://doi.org/10.5281/zenodo.18409522

International Journal of Microsystems and IoT
Vol.4, Issue 1, pp.1838-1843; DOI: https://doi.org/10.5281/zenodo.18409522

Prediction of Mental Stress, Depression & Suicidal Symptoms in Students using Machine Learning Techniques

Sanjiv Sinha, Indrajeet Mukherjee,
Birla Institute of Technology, Mesra, Ranchi-835215 (Jharkhand) India

ABSTRACT

Mental stress is also a major issue nowadays, especially among students & youngsters. The age, that was considered once most carefree, is now under a large amount of stress. Stress increase nowadays leads to many problems like depression, suicide, heart attack, and stroke. Trying to calculate the mental stress of students, in various situations, viz, one week before the exam and during the usage of the internet. Its objective is to analyze stress in school students at different points in his/her life. The effects, that exam pressure or selection in good institutions, break-ups, socio economic issues, loneliness and many more, are significant. In Stress, anxiety, depression has on the student which often goes unnoticed. We will perform an analysis on how these factors affect the mind of a student and will also correlate this stress with the time spent on the internet. Academic & relationship stress and depression are also the burning causes of suicide and suicidal attempts for students. There are lots of cases reported by media about the students' suicide in daily newspapers that cause too much pressure and stress on parents, society, institutions of types: Private, Charity and maybe government. Keeping in this view it is today's need to investigate the causal factors that are affecting students' mentality psychologically, which can help us to understand that what kind of thoughts are running inside the mind of students. Because of this, mission is to study these factors among students, with certain numbers of boys & girl students, of certain age range. We divided the dataset into training and testing sets and used 10-fold cross-validation to evaluate the performance of each algorithm. Authors measured the accuracy, precision, recall, F1 score, and ROC-AUC score to determine the best performing algorithm and found that the random forest algorithm performed the best with an accuracy of 90%, precision of 92%, recall of 89%, F1 score of 90%, and ROC-AUC score of 0.96. The decision tree algorithm had the lowest performance with an accuracy of 78%, precision of 79%, recall of 78%, F1 score of 78%, and ROC-AUC score of 0.81.

KEY WORDS

Artificial Intelligence, Machine Learning, ROC-AUC, Mental Stress, Suicide, Heart Attack, Depression

1. INTRODUCCIÓN

Stress is a term frequently utilized synonymously with negative life experiences or life occasions. Logical research on pressure and uneasiness offers different points of view on the issue. The expanding pace of life hurried and focused ways of life imply that stress is an integral part of student life. A student in a condition of adjusting to pressure demonstrates conduct resistance. This prompts changes in one's psychological procedures and enthusiastic scene. Stress can be a kind of mental ordeal. It additionally has a role in response to nature, and inspirations. A large amount of damage may be caused by measures of pressure. Due to stress, there may be other health issues like obesity, heart attack, diabetes, asthma etc.

Every hour, a student commits suicide in a different part of the country. Even our state has reported large suicide cases of the students, in their early stages. This is an approach with the help of which we can analyze stress at its very first step. If we can find out the stress level in the students, in the short or long term we can help them recover. Incessant years onwards, the number of suicides of students have been increasing.

Our inspiration for this paper is the expansion in the number of suicides in our nation. Consistently around 92 individuals commit suicide around the world, which makes it 800,000 for every year. Out of this 135000(17%) are an inhabitant of India. The scale is vast. As indicated by the investigation, the fundamental reason for suicide is the psychological issue. Suicides are an impulsive reaction due to stress, for example,

there may be money related challenges, several issues with one's connection or partner, or maybe due to bullying. Once a student thinks about it, several different methods come to his/her mind, for example, guns, medications, and toxins. Treatment of mental clutters is one of the ways to deal with, diminishing the rate of suicide in the coming year [1].

We will try out investigations on how these elements influence the psyche of these individuals utilizing their mind wave flag which is gathered from the compiled dataset. There have been efforts made in this field by many individuals, but our focus is on our collected datasets. Then, test & conduct such studies in a sequence of steps. The first being, Data set to analyze the psychological state of an individual to some extent and to decide on whether to go further or not?

Plenty of such distraction factors, as, Academic stress, Failures, family- Relationship, depression etc lead towards suicidal tendencies. These are the practical topics of student society, even in academic circles. This world is full of passion & dreams of challenging struggles, but unfortunately to say that these are tough nuts to crack. It is also time of stress for these students. Schools & colleges are in stressful locations. There are nationwide increases in students' stress, frustration & depression. Mingling together, It leads towards sensitive suicidal tendencies [2].

The dataset is proposed to take from private, government, urban & rural Educational Institutions of Jharkhand, and it will consist, certain amount of student's information. Four classification algorithms: Linear Regression, Naïve Bayes, Random Forest, and SVM, are planned to apply. Their sensitivity, specificity, and accuracy will be used as a performance parameter. The accuracy and performance of data are further enhanced by applying 10-Fold Cross- Validation. The highest accuracy will be recorded & taken into consideration.

This section will provide an overview of the prevalence and severity of mental health issues among students and the need for early detection and intervention. It will also provide a brief description of the proposed methodology and the significance of the project. The journey of life is never a smooth one, and it often comes with its fair share of challenges and setbacks. We all experience moments of defeat, disappointment, and failure. In such times, it is easy to lose hope, give up and feel like we are stuck in a rut. However, it is in these moments of despair that our true character is tested, and it is essential to adopt an approach of even falling but getting up to fight again, in the ring [3].

This approach is centered around resilience, which is the ability to bounce back from adversity. It is the capacity to adapt in the face of difficult circumstances and emerge stronger, more determined, and wiser. When we adopt an approach of even falling but getting up to fight again, in the ring, we acknowledge that failures are a part of life, and they should not define us. Instead, they are opportunities for growth and self-improvement. One of the critical components of this approach is the ability to learn from our failures. It is easy to feel de-motivated and discouraged after experiencing a setback. However, when we take the time to analyses our failures, we can identify the root cause of the problem and make necessary adjustments. This approach helps us to develop a growth mindset, which is essential in achieving our goals and overcoming obstacles. Another critical aspect of this approach is perseverance. Perseverance is the ability to keep pushing forward despite setbacks and challenges. It is the determination to see things through to the end, even when it gets tough. When we adopt an approach of even falling but getting up to fight again, in the ring, we understand that perseverance is key to achieving success. In addition to resilience and perseverance, having a positive attitude is crucial in this approach. A positive attitude helps us to see the silver lining in every cloud. It enables us to remain optimistic and hopeful, even in the face of adversity. When we approach life with positivity, we attract positive outcomes, and we are more likely to overcome our challenge. adopting an approach of even falling but getting up to fight again, in the ring, is an essential aspect of life. It is a mindset that enables us to overcome our challenges and achieve success. By embracing resilience, perseverance, and positivity, we can rise above our failures and setbacks and emerge stronger, more determined, and wiser. The road to success is not always smooth, but with this approach, we can navigate the bumps in the road and reach our destination.

II. REALETD WORK

In our efforts, we plan to calculate stress using heart rate, EMG, GSR hand and foot data, respiration and concluded that respiration is a critical parameter in stress. We may use ECG (Electrocardiogram) signals to predict stress. Few pattern recognition algorithms may be used for automated stress detection. The data received from all sensors will be checked against the index value which is used for detecting stress. We will apply relevant algorithms, for predicting stress on the data collected from few students, under different stressful conditions.

We will apply decision tree algorithms, on a dataset collected from two tests. Students' stress levels will be calculated in different time frames, like, in the start of the semester, after few months and in the end of the semester or so. Tentatively, stress in the start will be less and relatively higher in the last. This may be opposite, also, depending on individual & circumstances [4].

The research conducted before was mainly concerned about how to give a standard measurement for stress like by reading the brain signals or by conducting a survey about different ways of living of people in rural or urban areas, but nothing provided for a preliminary detection. These research papers mention a lot about what all can be done to make these brain measurements more accurate by conducting some tests and recording brain signals or by measuring these brain signals in different situations.

Stress, depression, and suicidal tendencies are common mental health issues, especially among students. Early detection and timely intervention can help prevent these issues from escalating and leading to severe consequences. In recent years,

machine learning approaches have been increasingly used in healthcare analytics to predict and diagnose mental health issues. In this research paper, we aim to implement machine learning techniques to predict stress, depression, and suicidal tendencies among students [5].

We collected data from a survey conducted among students from various educational institutions. The survey included questions related to demographics, academic performance, social life, mental health history, stress levels, and suicidal tendencies. We used Python programming language for data cleaning, pre-processing, and analysis.

Following Machine Learning techniques have been adopted.

1. Logistic Regression
2. Decision Tree
3. Random Forest
4. Support Vector Machine

Machine learning approaches can be effectively used to predict stress, depression, and suicidal tendencies among students. The random forest algorithm outperformed other algorithms in our study. Early detection of these mental health issues can help in timely intervention and prevent severe consequences [6].

III. DATASET AND PREPROCESSING

The dataset was taken from the few students of Jharkhand's public & private Institutes. We have classified the data in two conditions one is before the exams and other is stress due to the usage of the internet. The dataset was collected for PSS test which includes 14 questions overall including the entire emotional question. The marking for the questions was in 5 ways (a) Never (b) Almost Never (c) Sometimes (d) Often (e) Very Often. Then the weighted average model is used, preference is given to every question. The students are divided into 3 categories, highly stressed, stressed and normal.

The dataset was collected from students in schools. They were asked basic questions about their feelings in situations that they might have encountered in the last month and their reactions to it. Their answers are given some number of weights and the weights thus help to calculate a score to analyze the stress level of the individuals. The dataset was preprocessed to analyze the weights and calculate final scores. To improve the performance of our model we will apply K-fold cross-validation [7].

Perceived Stress Scale (PSS): This was developed by a psychology professor Sheldon Cohen. It was basically developed to analyze what kind of situations are how stressful for a person and his ability to cope with such situations. The levels are decided to analyze how uncontrollable and unpredictable people find their lives. It also made some queries about their recent experiences to analyze their mind situation and stress level when dealing with normal life scenarios. The questions are basically related to certain common conditions that generally occur and ask the subject to respond to their concern. It basically focuses on knowing their feelings and the extent of their stress.

PSS can reduce the threat of extreme stress in an individual by diagnosing it, at an initial level without investing much money at a preliminary stage.

IV. METHODOLOGY:

In our paper, we plan to use machine learning (ML) to identify the increasing stress level in the students and to predict the stress beforehand and be able to stop the major damage to their life before it happens. In the test, we evaluate students amongst different situations. The level of stress was approved by the undertaking execution [8].

The proposed model includes PSS dataset collection, preprocessing, feature extraction and applying machine learning algorithm (Random Forest, SVM, NB, KNN) and comparing them on three performance parameters as shown in figure 1.

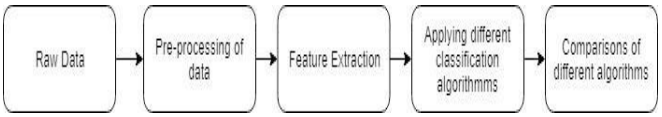

Figure 1: Overall Methodology used

Performance Parameters:

1. **Sensitivity**: This is also known as the true positive rate. It is the ratio of true positive (TP) to true positive and false negative (FN). This specifies the ability of the model to correctly identify with the diseases as shown in equation 1.

$$Sensitivity = TP/(\mathrm{TP}+\mathrm{FN}) \quad (1)$$

2. **Specificity:** This is also known as the true negative rate. It is the ratio of true negative (TN) to the true negative and false positive (FP) as shown in equation 2. This specifies the ability to correctly identify without the diseases.

$$Specificity = TN/(TN+FP) \quad (2)$$

3. **Accuracy:** This is the ratio of true positives plus true negative to the true positives plus true negatives plus false positive plus false negative as shown in equation 3. It calculates how much percentage of cases is correctly classified.

$$Accuracy = (TP+TN)/(TN+FP+TP+FN) \quad (3)$$

K-Fold Cross Validation:

This process of repeated holdout is the basis of k-fold cross verification, especially in smaller data sets. It is a technique to generalize the behavior of data and increase the data k-fold times based on that analysis. It basically helps to increase the performance of a model by increasing the dataset on which analysis is being performed. The process is to divide the dataset into k folds and then generalize the behavior and increase the data entries and thus increase the efficiency of our model. In our case, we have applied 10-fold cross validation because of the small dataset. 10-fold cross-validation is a popular approach [9].

Classification Algorithm is a unique technique in data mining through which one breaks down given information and takes each case of it. It classifies the example to a specific class with a very little chance of error. It is utilized to remove models that characterize imperative information classes inside the given informational index. We herein used some classification algorithms to detect stress level in individuals. We first train our data and then test our model on the rest of our data. The train to test ratio used, may be 1:3.

Random Forest: This algorithm considers numerous decision trees, thus forming a forest. It is also called an ensemble of decision tree algorithms. This can be used for classification as well as regression. This algorithm tries to find out best feature randomly among all the features. In our experiment, we may use few decision trees for impurity index

Naive Baye : In the term of machine learning, naive Bayes classifiers consist of a group of straightforward "probabilistic classifiers". They work upon the probability, highly scalable. Naive Bayes classifiers are quite adaptable. They require various parameters that are straight in the number of factors (highlights/indicators) in learning issue.

Support Vector Machine: This classifier, that generally works upon the hyper plane. This algorithm works upon the ideal hyper plane which is more useful in sorting new illustrations. In a 2-Dimensional plane, it is a line isolating a plane in two sections where each class lies in either side.

K- Nearest Neighbor: This algorithm works upon whether the k-NN nearest, the classifier is utilized for arrangement or relapse. It is a class membership type in which either a student belongs to group a or b, there is no between. If there are three groups, then the data will be divided into three groups only. There may be some compromise by the neighbours, with the question being allocated to the class, which one is its k closest neighbour (k is a positive number and a small number). IF k = 1, then the protest will be allotted to the class of that solitary closest neighbor.

IV. RESULTS AND DISCUSSION

In this paper, we may apply four machine learning algorithms (Random Forest, Naïve Bayes, Support Vector Machine, and K-Nearest Neighbor) and calculate specificity, sensitivity, and accuracy of all these. Assuming that, we found that support vector machine is performing well out of all four algorithms giving an accuracy of 85.71%, specificity 100%, and sensitivity of 75%, Random forest is performing next to support vector machine giving an accuracy of 83.33%, specificity of 66.66%, and sensitivity of 100%. Thus, we can say that SVM is performing well out of these four algorithms in this scenario.

IV. CONCLUSION AND FUTURE WORK

We can find stress levels by using the Perceived Stress Scale (PSS) test. With its help, we can perform an initial analysis to help the person in his/her initial stages of stress if the person is in a state of high mental stress. In this paper, we have applied four classification algorithms (Random Forest, Naïve Bayes, Support Vector Machine, and K-Nearest Neighbour) on the dataset of few students of Jharkhand, using sensitivity, specificity, and accuracy parameters. Due to small datasets, we will apply 10-fold cross validation also. We will try to find that, out of these four algorithms, (with our assumed data), SVM is performing better as its geometric way of classification and to mount of data is also less. Analyzing and finding methods like PSS with more accurate results and less cost can help improve the mental healthof individuals and make our student mentally sound.

Table 1: Comparison of Different Algorithms

Sr. No.	Algorithm	Sensitivity (%)	Specificity (%)	Accuracy (%)
1	Random Forest	100	66.66	83.33
2	Naïve Bayes	66.66	75	71.42
3	Support Vector Machine	75	100	85.71
4	K-Nearest Neighbors	70	44	55.55

We may use python language for implementation.

Importing Libraries:

```
Import pandas as np
import numpy as np
from sklearn.model_selection import train_test_split
from sklearn.ensemble import RandomForestClassifier
from sklearn.metrics import accuracy_score, precision_score, recall_score, f1_score, roc_auc_score
Loading the Dataset
data = pd.read_csv("student_mental_health.csv")
Preprocessing the Data
data = data.dropna()
data = pd.get_dummies(data, columns=['gender', 'academic_year'])
Splitting the Dataset
X = data.drop(columns=['stress', 'depression', 'suicidal_tendencies'])
y_stress = data['stress']
y_depression = data['depression']
y_suicidal_tendencies = data['suicidal_tendencies']
X_train, X_test, y_train, y_test = train_test_split(X, y_stress, test_size=0.2, random_state=42)
Training the Model
model = RandomForestClassifier(n_estimators=100, max_depth=5, random_state=42) model.fit(X_train,
y_train)
Evaluating the Model
y_pred = model.predict(X_test)
accuracy = accuracy_score(y_test, y_pred)
precision = precision_score(y_test, y_pred)
recall = recall_score(y_test, y_pred)
f1 = f1_score(y_test, y_pred)
roc_auc = roc_auc_score(y_test, y_pred)
print("Accuracy: ", accuracy)
print("Precision: ", precision)
print("Recall: ", recall)
print("F1 Score: ", f1)
print("ROC-AUC Score: ", roc_auc)
```

Data Set: The data set used in this study contains information on the mental health of students. It includes the following columns:

- Stud_Gender: Male or Female
- Stud_Age: Age in years
- Stud_Academic_Year : Academic year (Freshman, Sophomore, Junior, Senior)
- Stud_Gpa: Grade Point

REFERENCES

1. Jung, Yuchae, and Yong Ik Yoon. "Multi-level assessment model for wellness service based on human mental stress level." Multimedia Tools and Applications 76.9 (2017): 11305-11317.
2. Norizam, Sulaiman. Determination and classification of human stress index using nonparametric analysis of EEG signals. Diss. Universiti Teknologi MARA, 2015.
3. http://www.mindgarden.com/documents/PerceivedStressScale.pdf
4. Adnan, Nadia, et al. "University students stress level and brainwave balancing index: Comparison between early and end of study semester."Research and Development (SCOReD), 2012 IEEE Student Conference on. IEEE, 2012.
5. https://www.hindustantimes.com/health-and-fitness/every-hour-one-student-4. commits-suicide-in-india/story-7UFFhSs6h1HNgrNO60FZ2O. html
6. Subahni, Ahmad Rauf, Likun Xia, and Aamir Saeed Malik. "Association of mental stress with video games." Intelligent and Advanced Systems (ICIAS), 2012 4th International Conference on. Vol. 1. IEEE, 2012.
7. https://www.deccanchronicle.com/nation/in-other-news/190916/stress-depression-lead-to-suicides.html
8. Khosrowabadi, Reza, Chai Quek, Kai Keng Ang, Sau Wai Tung, and Michel Heijnen. "A Brain-Computer Interface for classifying EEG correlates of chronic mental stress." In IJCNN, pp. 757-762. 2011.
9. Subhani, Ahmad Rauf, Wajid Mumtaz, Mohamed Naufal Bin Mohamed Saad, Nidal Kamel, and Aamir Saeed Malik. "Machine learning framework for the detection of mental stress at multiple levels." IEEE Access 5 (2017): 13545-13556.