Projective Exploration on Individual Stress Levels Using Machine Learning

Nagendra Prabhu S
Department Of Computer Science and Engineering Department,
New Horizon College Of Engineering
Bangalore,India
snagendraprabhu@gmail.com

Dr. Amarjeet Singh
Dean Academic,
Department Of Computer Science and Engineering Department,
New Horizon College of Engineering
Bangalore,India

Bhavya B
Department Of Computer Science And Engineering Department,
New Horizon College Of Engineering
Bangalore,India
bhavyab0203@gmail.com

K. Sreeja
Department Of Computer Science And Engineering Department,
New Horizon College Of Engineering
Bangalore,India
k.sreeja2001@gmail.com, @gmail.com

Bhargavi P
Department Of Computer Science And Engineering Department,
New Horizon College Of Engineering
Bangalore,India
p.bhargavi28112811@gmail.com

***Abstract*—Recently, Stress Prediction in every individual based on their profile and behaviour is a challenging task in the current sector. Current system is a manual process where it is difficult to identify the stress in the college students or employees. There is no automation for any stress prediction. System uses machine learning algorithms or AI algorithms to find out the stress levels, these technologies are used for application development. The purpose of this project is to reduce stress in students and employees. Computer science has come a long way in the last few years. It is massive and multifaceted. It has been used in a variety of applications to meet the basic needs of human society. In the field of healthcare, machine learning has made significant progress. Stress is a deadly disease that kills a large number of people worldwide. We examine how machine learning techniques can assist reduce the risk of stress prediction, which can lead to accidents, in this study.**

Keywords: Machine Learning, ASP.NET, Stress detection, social communication, Labelled and unlabelled data set.

1. INTRODUCTION

In the understudies of students such as graduate, undergraduate, postgraduate, and professional students, machine learning allows for the prediction of the potential of stress prediction. We examine the effectiveness of machine learning algorithms in reducing the risk of stress prediction, leading in early treatment of under studies students, in this work. Nowadays, stress became a major problem and sometimes it may response in positively or negatively. Now a days the word stress became a major thing which leads to different health problems when it crosses certain level, it complexes day by day.

Now a days both females and males can be introverts. Introverts suffer because they don't understand themselves and they aren't understood when surrounded by extroverts.

So, this project helps them, an introvert may believe that going through an experience in great detail will eventually result in figuring out what went wrong according to those suggestions will be given and help them from stress.

Fig .1 Indicates Stress

In today's culture, the number of people suffering from stress is steadily increasing. The effects of stress on one's entire health are intertwined. According to Hans Selye, stress is "an organism's non-specific response to a demand or a change in its physical circumstances." It might manifest as either eustress or distress. Eustress has a beneficial effect, whereas distress has a negative effect. A pattern formed by cognitive, emotional, behavioral, and physiological responses to negative and toxic characteristics of a stressed organization and environment. It's the outcome of misalignment between people and their jobs, as well as interpersonal friction.

This will assist students, employees, and all individuals in objectively determining the correct vision for themselves. From a psychological standpoint, self-reported measurements are the most common way to assess stress. According to the literature review, data mining tools have not

yet been widely used to analyses stress issues; thereby, we use machine learning techniques to overcome this stress issue.

A. Popular Algorithms

- KNN Classifier: It is a supervised learning algorithm which can be implemented on labelled data. It is used for predicting whether a person needs treatment or not. From the data already known it classifies on how similar its independent variable are to the instance with dependent
- Logistic Regression: The best method that is used for prediction analysis. It is used with binary variable that is dependent on other independent variables. Here, we will use few relevant attributes as independent variable and those having stress and needs treatment that is predicted by the model trained.
- Random Forest Classifier: It is used as the flexible ML algorithm which provides good persistent result even if there is no hyper tuning. They are used as the clusters of decision trees that are working together.

Input Dataset $D = D^S \cup D^T$ with feature representation $\boldsymbol{x}_i$

Output Dataset $D = D^S \cup D^T$ with feature representation $\boldsymbol{f}_i$

1. Train the LLP classifier with D
2. Predict the pseudo-label values for target data
3. Noisy labels dataset: $P = \{\boldsymbol{x}_i, y_i\}_{i=1}^{n_S+n_T}$
4. **for** r =1 to R **do**
5. Training subset: P^r={m random samples}
6. Train projection functions $f^r(.)$ with P^r using eqn. (4)
7. **end**
8. **for** i=1 to $n_s + n_T$ **do**
9. **for** r=1 to R **do**
10. Obtain projection vector: $V_i^r = f^r(\boldsymbol{x}_i)$
11. **end**
12. $V_i = \left(\left(V_i^1\right)', \ldots, \left(V_i^R\right)'\right)'$
13. **End**

Fig .2. pseudocode for Logistic Regression

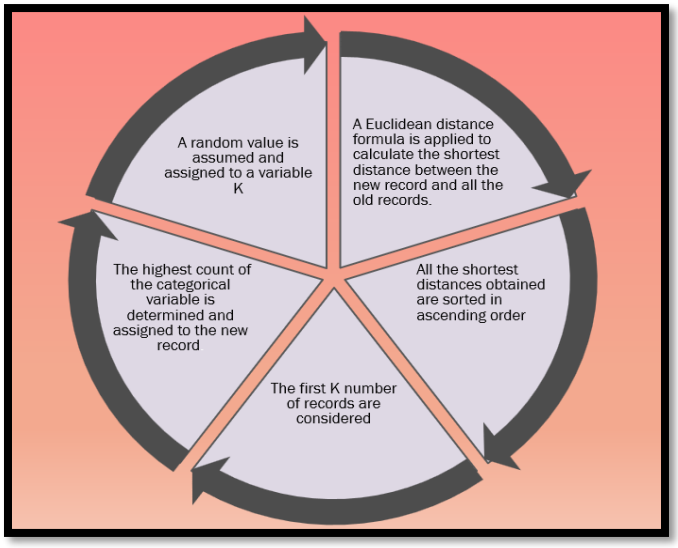

Fig.2. .Working of KNN algorithm

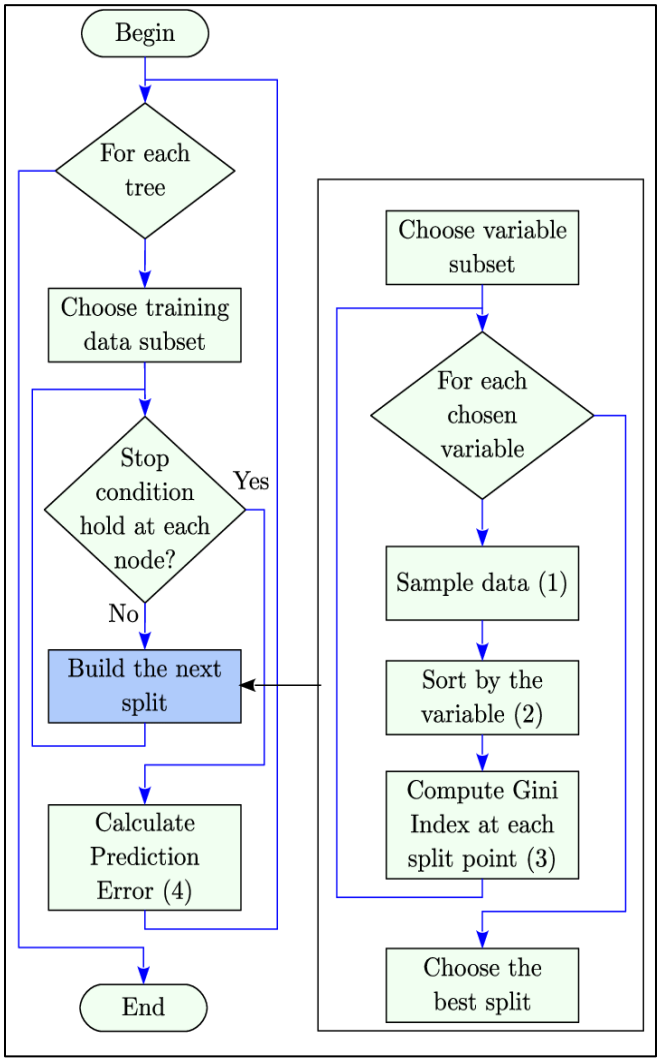

Fig .4.Random Forest algorithm

II. STRESS TYPES,CAUSE AND SYMPTOMS.

Hypothyroidism/underactive thyroid is a disorder or the condition of the endocrine system in which the thyroid gland does not produce enough thyroid hormone. A very less quantity of iodine in the diet is the major cause of hypothyroidism. Hashimoto's thyroiditis is the most common cause in countries with sufficient dietary iodine. When suspected, can be confirmed with blood diagnoses measuring thyroid-stimulating hormone (TSH) and thyroxine levels.

A. Potentially harmful factors

Any stimulus that causes the release of stress hormones is referred to as a stressor. The following are the two most prevalent forms of stressors:

1) Physiological (or bodily) stresses can include high temperatures, physical injury, pain, pollution, infection, and a lack of relaxation, to name a few.

2) Negatives include personalities events such as death of a person, new/fired from work, missing deadlines, the changes in social life or sleep habit, money troubles, and multitasking job can all induce psychological pressures.

B. Workplace Stress

When job demands exceed an employee's skills, it raises the chance of workplace stress and lowers productivity. Workload, a noisy atmosphere, frustrated dealers, work pressure, low performance review, and supervisor-worker interactions are all examples of workplace stressors. As a result, some employees increase their smoking environments, become unhappy, taking from others, losing appetite, and perform poorly. Individuals face stress when they have least control over the quantity of duties they are allotted, according to a study result. According to statistics, 41% of American adults are stressed by financial concerns. Employees are not present in Australia.

C. Types of Stress

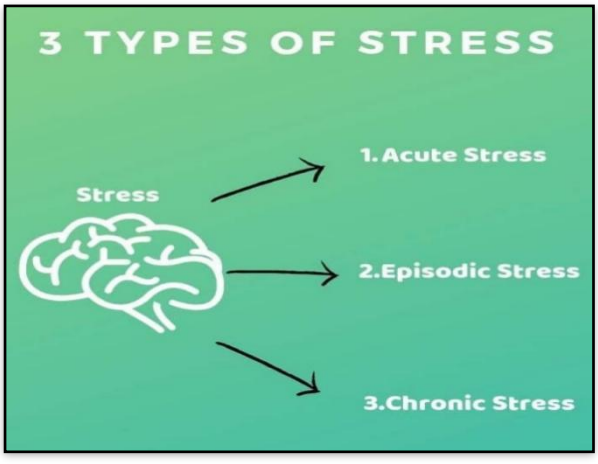

Fig .5 Stress types

1. Acute Stress:
   - Definition: It is a type of short-term stress that comes as a result of daily duties, recent events, or events in the near future. It causes emotional suffering (irritability, anxiety, and anger) as well as medical difficulties.
   - Cause:
     - Witnessing a death.
     - Witnessing a severe accident.
     - Exposure to war in the case of either civilians
     - Rape or sexual violence.
     - A physical attack.
   - Symptoms:
     - Emotional responsiveness.
     - Difficulty concentrating.
     - Feeling of being disconnected from one's body
     - Sense of numbing
     - Dissociative amnesia.

2. Episodic Acute Stress:
   - Definition: It occurs when people are confronted with a series of stressful situations and are aware of impending danger. Aggressive, uptight, and extremely competitive people with a "Form A" mentality are prone to this type of stress. Continuous anxiety, migraines, heart/chest pain, and persistent headaches are all prevalent signs of episodic stress.
   - Cause:
     - Diabetes
     - Depression leading to suicidal tendence.
     - Menstrual and skin related problems.
     - Loss of memory
   - Symptoms:
     - Panic attacks
     - Irritability
     - Muscle tension

3.Chronic Stress:
   - Definition: It is the most dangerous type of stress. It entails constant anxieties that appear to last forever, affecting both the mind and the body.
   - Cause:
     - Acne
     - Depression
     - High blood pressure
     - Heart disease
     - Eczema
   - Symptoms:
     - Aches and pains
     - Difficulty sleeping
     - Fatigue,Infections
     - Feeling a loss of control

III. LITERATURE SURVEY

Authors **Disha Sharma et al. (2020) [1]** describe that stress among the students they collected the information provided by the students in the form of a quiz or self-rating called a dataset and they used this dataset to make an effective analysis on the stress levels of every student. The information for getting a dataset were collected using both primary as well as secondary fashion. The survey was conducted in the following order such as E-mails, google Drive, survey forms etc., for each questionnaire they provided the students with 5 choices to choose one among them and rate themselves in opinion to their stressful situation. Choice includes

(A) Never
(B) Almost
(C) often
(D) Very Often
(E) Sometimes

Which had its own rating rank so that it would be easy for detecting the stress faced.

Therefore, the data that was collected is being put together and made use on an pss(perceived stress scale). By using this source, they discovered the usage of parameters such as True positive rate, false positive rate, recall and f-score.

The formulas for the above is as follows:

TP=True positive, FN=False negative, P=precision.

- TP = TP+ TP/TP +FN.
- FP = FP + FP/TN.
- F-score=2*P*recall/P+ recall.

As per this research we found that they used the technologies such as Naive bayes; Random Forest; Bayes' net; logistic regression.

On comparing, the tool used was weka tool. The overall results on considering various techniques, calculations and comparing we get the following result that is Bayes' net classifier gives the highest accuracy amongst all the other algorithms and provided the final accuracy as 88.5965% as an overall percentage for a particular student's stress level.

Authors **Jarernsri Mitrpanont et al. (2017) [2]** presents a method to create Emotive Device which defines BCI (Brain Computer Interface) technology is used to assess the values that are sensed by sensors which are placed on human scalp. The basic health Questions are a set of 30 questions used to assess stress levels.

The features of this project is:

- Device
- Detect Stress
- Intervention
- Result,
- Observation data report.

Android 4.4, SQLITE, Adobe photoshop, Bluetooth to connect with emotive device, Java programming language these are the following techniques used in this project. Thus, concentration, relaxation and therapy of music interventions are all included in the proportion of stress. To summarise, this paper it found to be a very important tool for assisting individuals in coping with their own stress by utilising the most successful approach known as "Neurofeedback," which gives a perfect option in treating and matching with each and every individual.

Authors **Enrique gorcia et al. (2015) [3]** compares the methods to Auto detect stress among the working environments, they collected the information based on behavioural data which was provided by using the built-in sensors, especially with the use of Samsung galaxy STTI mini smartphones. This dataset was collected in the form of smartphones, written format using applications. From the survey conducted for each questionnaire which resulted in the form of low, medium and high, also it is found that it can identify the rate of stress levels day by day with the help of a smartphone. In accordance with this paper, we found that the technologies used were Naive bayes; and decision trees. By considering various techniques their experiment included 19subjects. Thus, the overall accuracy was found to be 75% which was the first thing on evaluating the potential from mobile phones as a stress detector in the in environment.

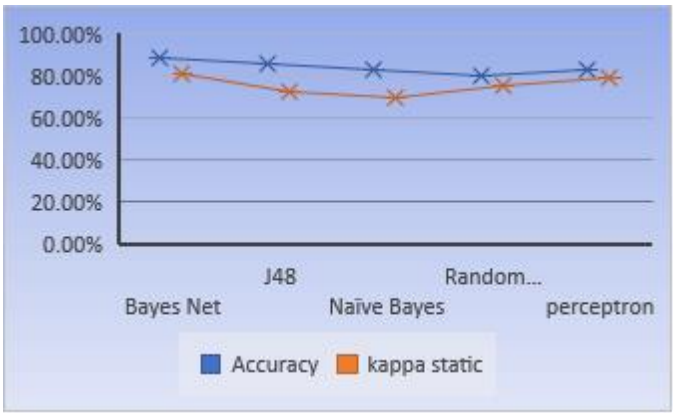

Fig .6 Algorithms based grapgh

Authors **MS. Sumathi M. R and Dr.B. Poorna (2016) [4]** proposed the method to predict attention problems, academic problems, anxiety problems, POD.

In order to predict this, they used Weka tool and artificial intelligent techniques such as Aodesr, multilayer perceptron (MLP), lestar, multi-class classifier (MCC), Ft and Lad tree. This collected information is formed as a dataset with 60 instances.

Therefore, the overall accuracy classifies into different areas such as Kappa statistics, accuracy, and loc area, on comparison among all the attributes and the selected attributes the highest accuracy was found in multilayer perceptron, multi-class classifier and LAD tree.

Authors **Xiyu Liu et al. (2020) [5]** demonstrates model predict the stress among the college students they collected the information provided by the students in the form of a questionnaire which is designed in the form of four dimensions:

- General Demographic
- Epidemic presentation
- Control in the region
- Psychological stress

For each questionnaire they provided the students with 4 choices to choose one among them and rate themselves in opinion to their stressful situation. Choice includes

A. All the time
B. None

C. Occasionally
D. Not clear

They used spss statistics 25.0 software for calculating accuracy. By calculating stress level using above choices it results the emotional display like worry, tension, anxiety, depressed, fear and mood swings. Thus, they can measure of mental health of a student in a covid pandemic through this project.

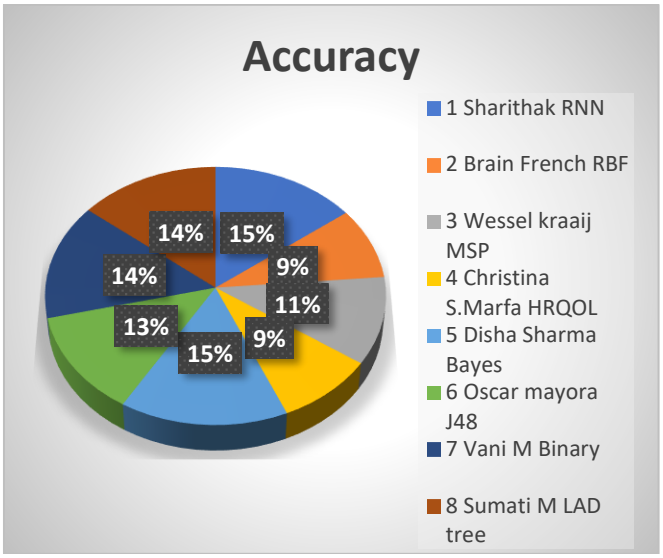

Fig .7 Accuracy differentiation

Authors **Yuan Shi et al. (2019) [6]** described We build a personalized stress detecting model based on SVM, and examined it on the collected data. We examine topics like continuous data, non intrusive detection of physiological measures, involving data collection, feature extraction, and model construction. The concepts which was ECG, GSR, RIP, EMA, RBF kernels.

- Sessions in the lab study.
- EMA questions and possible answers.
- Features extracted from sensor data
- Numbers of +ve and -ve from subjects

Average and standard errors of precision values
The final result the precision–recall curve of the stress detection model with personalization, segment-based features, and RBF kernel. Even at 100% recall, the model can still achieve high precision over 50%.

Authors **Saskia Koldijk et al. (2021) [7]** compares the Employees often record their experience of stress at work.
This is a literature survey on the paper which is done based on ML, mental stress inference, computer log, detection of facial expressions, gestures and posture.

- There are many options on how to deal with the aspect of time.
- There is debate on whether subjective ratings provide a good ground for truth.

The final performance of Accuracy is 0.8416 which has been reached.

Authors **S. Malfa et al. (2021) [8]**
Suggests the health-related quality of life and psychological stress between public sector and professional groups.
1.Other main characteristics of the participants by this professional group include.
2.Anxiety,depression and HRQoL are the scores (means and SD) by professional category and between- group comparisons.

Authors **Simhadri Naga Mounika et al. (2019) [9]**
The purpose of this paper is that Stress plays a very crucial part in our daily life. This is a literature survey on the paper which is done based on the deep learning concepts those are CNN, RNN, DNN algorithms, FGM for the graph implementation model, SVM for the vector machine, and finally with the neural network. The accuracy is predicted based on the percentage of positive, negative and Neutral.
The highest accuracy is 81.6 for RNN and for CNN is 69.8.

Authors **Pramod Bobade, et al. (2020) [10]** demonstrates The task was divided into three categories: amusement, baseline, and stress using Dl and ml. The non-stress class was created by combining the amusement and baseline states. Second, the non-stress class was created by combining the amusement and baseline states.
In order to predict this, they used WESAD dataset study. This dataset includes Data from numerous physiological modalities such as 3-axis acceleration (ACC), respiration (RESP),
and volume blood pulse are also included in the WESAD dataset However during the overall analysis, machine learning techniques yielded accuracy of 81.65 % and 93.20 % for 3-class and binary classification problems and deep learning yielded accuracies up to 84.32 % and 95.21 %, and this was finally predicted as the overall outcome of this paper.

IV. METHODOLOGY

As the starting step for the research the problem is to identify the stress levels in every individual by conducting a survey through google forms or emails.
The collected data is now being understood and trying to fetch the required data for processing here the terms that have most similarities will be grouped based on choices. A model will be built using machine learning techniques and data accessing techniques such as ADO.NET.These will be further classified with the help of several algorithmic techniques, like for this approach we are trying to implement using KNN-classification algorithm and further is classified into sub divisions such as supervised learning and unsupervised.

Fig .8 System architecture

Attributes are identified by the dataset and listed in the form of a table. Where all attributes are of nominal types whereas some attributes which is relevant for classification and prediction for the problem, here the dataset will be pre-processed by eliminating irrelevant and redundant attributes using three tier architecture they are data layer, business layer and presentation layer.

EXISTING SYSTEM:

In the existing system college students are facing so many mental health problems such as depression, pressure, stress, interpersonal sensitivity, fear

Though many industries and corporations provide mental health related schemes and try to ease the workplace atmosphere, the issue is far from control.

LIMITATIONS OF THE EXISTING SYSTEM:

• No Automation for stress prediction
• Less user satisfaction
• Time consuming
• Less Reliable
• Less Efficient
• Manual Process

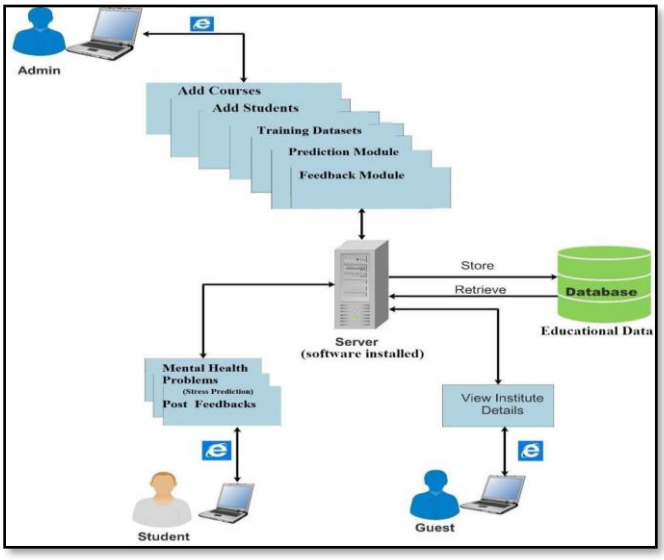

Fig .9 Overview of functionalities of the module

V.CONCLUSION

- This project helps in predicting the stress levels due to various parameters considered through a survey and it also lists as to whether the students are stress free or stressful and the range of their stress levels.
- The student can incorporate the solution and work towards maintaining his or her mental equilibrium.
- So, this project helps them, an introvert may believe that going through an experience in great detail will eventually result in figuring out what went wrong according to those suggestions will be given and help them from stress.
- Machine learning methods for prediction of stress and mental health condition gives significant results and can be explored further, meeting the objectives of this paper.

VI. FUTURE WORK

- With changing lifestyle and work culture, there is an increase in the risk of stress among the college students. System find factors that strongly determine the stress levels.
- Stress was identified based on gender, family history and availability of health benefits in this sector.
- By identifying the stress of among individuals, system comes up with some approaches to reduce stress and create a much comfortable zone.
- System uses many parameters such as gender, age, family history, provided health benefits, sharing about illness, tech company, tech role, acquiring leave etc.
- System uses machine learning algorithms or AI algorithms to find stress of a student.
- System can be developed as a real time application which is useful for colleges. As Visual Studio and SQL Server is more supportive with real time applications, these technologies are used for application dev

VII. REFERENCES

1. Disha sharma, nikitha Kapoor, Dr.sandeep. “Stress Prediction of students using machine learning”. Trans stellar, vol.10,issue 3, June 2020.
2. Jarernsri Mitrpanont, Jaruwan Phandhu-fung and Nantanut Klubdee.” iCare-Stress: Caring System for Stress”. European journal of operational research, 2017.
3. Enrique gorcia, venet osmani and Oscar mayora. “Automatic stress detection in working environments from smartphones”, 14 oct 2015.
4. MS. Sumathi M. R and Dr.B. Poorna. “Prediction of mental health problems among children using machine learning techniques”. International journal of advanced computer science, vol7, nov 1,2016.
5. Xiyu Liu and Yanliang zhang.” How COVID-19 affects mental health of college students and its countermeasures”.

International conference on public health and data science (ICPHDS),2020.
6. Yuan shi,Minh hoai Nguyen, Patrick belt.” Personalized Stress Detection from Physiological Measurements”. IEEE, 2019.
7. Saskia Koldijk, Mark A.” Detecting work stress in offices by combining unobtrusive sensors”. Radboud respository university, Nov 09,2021.
8. Christina S. Malfa,” Psychological Distress and Health-Related Quality of Life in Public Sector Personnel”. Environment research and public health, feb 14, 2021.
9. Simhadri Naga Mounika, Prem Kumar.” Detection of Stress Levels in Students using Social Media Feed”. IEEE, 2019.
10. Pramod Bobade, Vani M.” Stress Detection with Machine Learning and Deep Learning using Multimodal Physiological Data “IEEE, sep 06,2020
11. Rajeswari S,Kalaiselvi R. “Survey of data and storage security in cloud computing”. IEEE International conference on circuits and systems, ICCS 2017,2018
12. Prabhu Nagendra S, Shanthi D. “A survey on Anomoly Detection of Botnet in network”.International Journal of Advance Research in Computer science and management studies,Vol 2,Isuue 1,2014.