

# **STRESS LEVEL DETECTION**

Kommineni Venkata Naga Kaushik  
UG Student  
Department of Computer Science and Engineering  
KL University  
Hyderabad,India  
[kaushik.kommineni@gmail.com](mailto:kaushik.kommineni@gmail.com)

**Abstract** — Due to the COVID-19 pandemic, to control pandemic situations and its spread, the government took a decision to shut all the educational institutions, which in turn creating a direct impact on many people by causing stress and mental illness. We propose a solution for organizations where they can know the levels of stress faced by the students and could calculate percentage of stress. So for this to be done, students can take up the survey through a google form which consist of the parameters which are helpful in collecting information about mental distress and many other psychological factors faced by the students. The data which is collected from the students is inputted into the model with results the stress levels of the students.

**Keywords:** K-Means, Principal component analysis (PCA), Logistic Regression.

## I. INTRODUCTION

In a pandemic, the people's outlook of health-care constraints and lifestyles is completely switched. Since then, covid 19 had spread a lot, causing global disturbances.

The administration of educational institutions has closed across the globe to prune the growth of the disease and in welfare of all people.

Considering all these circumstances the people around all countries were affected by entities like food availability and medical facilities.

Many surveys were conducted to study the person's stress level based on the stress constraints like physiological conditions. A person can be stressed out in scenarios like worrying about losing their employment, family health conditions and about the grades in examinations. These kinds of stressful scenarios increase the stress levels which affects the increase of heart and muscles related issues.

Generally, anxiety and stress are very much common among all the students with a variation of degree. So, by observing each and every student it would be a huge task to go through their profiles. This problem makes us create a new model for predicting stress automatically. For each student who is undergoing various psychological parameters of stress and proposes a solution for that.

So for this to be done, some Machine Learning algorithms and Data Science techniques are used. Maintaining track records of each student's stress levels, and studying them makes us understand the degrees of stress of the students in organization .

Students are categorized into 2 sublevels in regards with the stress percentage they face : i.e.,

over-stressed or under-stressed. And according to that, the range of stress is highlighted based on the levels.. Based on this percentage, the authorities give advice to the students. As a result, we create a model for unlabeled data and untrained data that will determine the stress level of students using different Machine Learning and data science techniques.

## II. LITERATURE SURVEY

Paper [1], In software engineers, stress has been highlighted as a major issue. This dataset has been collected and submitted to a variety of machine learning algorithms .According to their research, 75% of employees working in the IT field are susceptible to pressure. Boosting and decision trees are just a few of the strategies they've worked on.

In Paper [2], The authors used many techniques to evaluate stress, including EEG, GSR, EMG, and SpO2. by these parameters stress levels of a person are calculated automatically.. Sensors are used here to collect the data from a person and then are compared to the indexed value that is given as input to create a baseline of comparison.

In Paper [3], The authors applied a decision tree algorithm to analyze data from students whose stress levels were measured at the start and end of the semester. As a consequence, the model identifies students who are more stressed near the conclusion of the semester rather than at the beginning.

In Paper[4], Distance education was made easier with higher development in technology with time. Many terms like Internet learning, web-based education, etc., have one thing in common i.e., stay connected with the computer through a series of networks . So by any means people are connected to the internet and learn from it such that people are distancing from each other and facing little mental illness.

In Paper [5], The author of this paper created a new model, which uses supervised algorithm labeling of data. Here supervised labeling method is used in contrast to an unsupervised clustering method which analyzes the ambiguous work and assigns or labels the specific work as per its senses.

In Paper [6], It stated that the increase in technology and creating new things in our

everyday life is taking an important part of our life. Day to day life innovations include all the new technologies.

### A. EXISTING SYSTEM

**Stress Detection Using Decision Tree and KNN:** In existing systems, the data set is collected from the online where, the data and some attributes like Heart rate of a person, Respiration rate, Electrocardiogram, Electromyography, Galvanic skin response, etc., are taken and this input is processed.

Based on the values, some of the algorithms such as Decision Tree, K-Nearest Neighbor are used to get accuracy. The dataset is compared with the threshold values and that value will compare and predict whether the person is in stress or not stress.

**Drawbacks of the Existing System:** The KNN algorithm is used in existing systems, but it does not work well with large dataset. So to reduce the dimension, additional work should be done before applying the KNN algorithm to any dataset. In contrary, wrong prediction are generated using this algorithm. This algorithm is sensitive to noisy data and missing values .

### B. PROPOSED SYSTEM

Detecting the stress level of students due to online education and prevention of stress of online education on students. This model does not illustrate any unique conceptual donation to the health care sector during the pandemic circumstances but provides the idea of improving the mental health of students .

We will use some data science clustering methods like K MEANS CLUSTERING to create preliminary clusters and then we will try to apply machine learning algorithms to classify students' stress levels during an online class.

**Advantages of the proposed system:** K-means algorithm is used in the proposed system , where it is suitable for larger dimensions of datasets. It is computed much faster than the smaller dataset. It can also produce higher clusters. The data which is raw is mostly an unlabeled one without any particular label or class defined with it. So it will be beneficial if we label the data using some new and efficient algorithms like K-means to group into clusters of

k numbers.

### III. DATA FLOW DIAGRAM

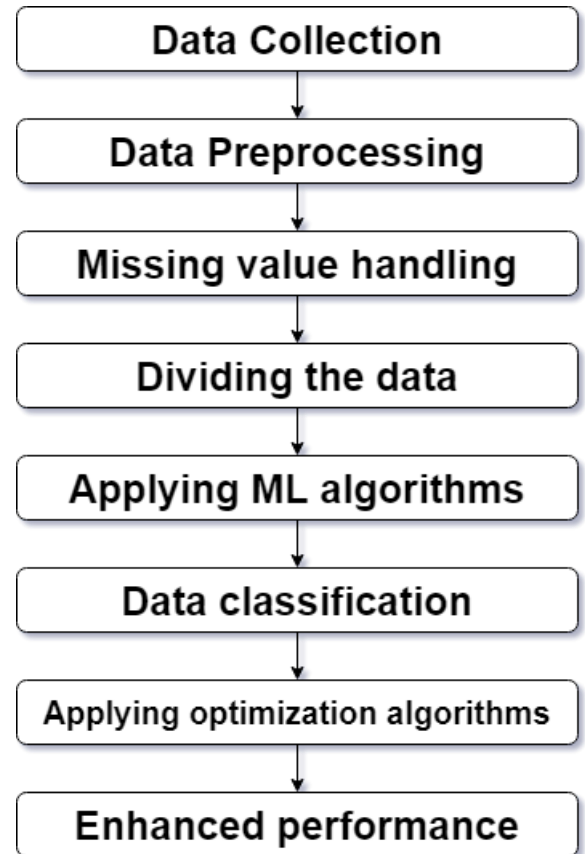


Fig1 . FlowChart Diagram.

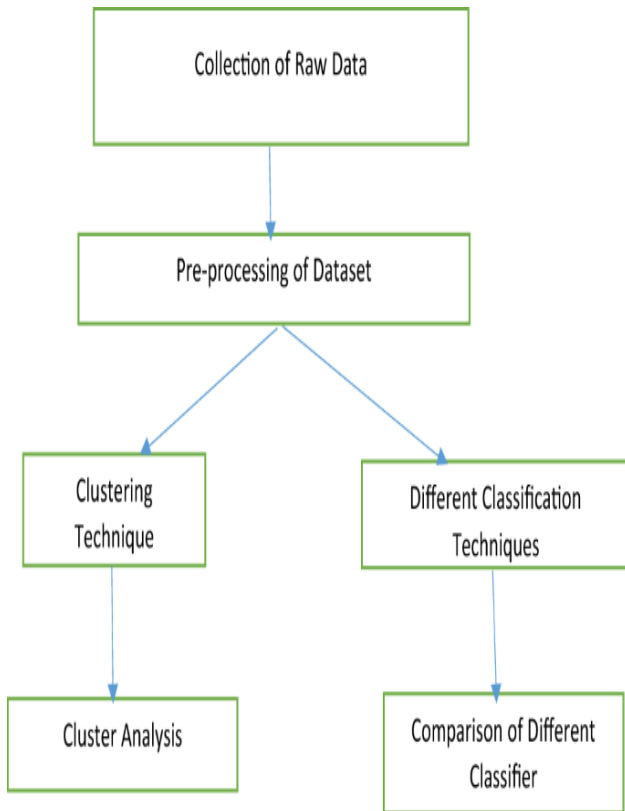


fig 2. Process of proposed system

**Data Collection:** Initially data is collected from the students using a google form by asking a series of questions, and replies of students are saved in excel.

**Data Processing:** And then processing of data takes place, where data is encoded together in a machine understandable way.

**Hiding Missing Values:** As ML does not support missing values, handling them will be the next step.

**Dividing the data:** Then next the data collected gets separated into training and testing sets. So that would be easier to apply ML algorithm.

**Applying ML Algorithms:** Now ML algorithms are added which allows the project outcomes to be more accurate and precise

**Data Classification:** As there is a vast amount of data, Data classification takes place to label the output classes.

**Data Optimization:** After Data Classification, Data Optimisation should be done for the trained data set. This is very useful

because we can bet on the best design with the required priorities . And also we could get results error free.

We'll go through all the steps with our dataset as we mentioned above. It is essential to go through all these steps to achieve our goal to find the stress level of students due to online education.

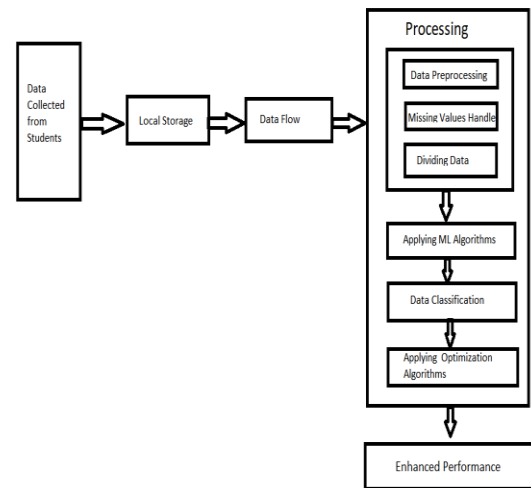


fig 3. Data flow diagram

#### IV. MODULES

**User Module:** Firstly, Users will receive a google form and then they will fill the google form.

**Admin Module:** We'll collect the data from students through their feedback filled in the google form and that will be stored in an excel sheet. Considering the feedback from 500 students, the dataset was constructed.

Uploading the file : Upload that csv file into a jupyter notebook.

**Processing Module:** Packages used here are pandas, NumPy, sklearn, LabelEncoder, PCA, KMeans, LogisticRegression. Here checking of unique and null values takes place and we'll drop the null values. Then by using Label encoders, labeling the data is done.

Now, by considering the labeled data PCA minimizes the dimensionality of the data. By using k-means we'll predict the analysis by dividing the dataset into k clusters and then we'll label the unsupervised data. And finally we'll plot the labeled data of stress levels of the students.

: <matplotlib.collections.PathCollection at 0x205e3d14ac0>

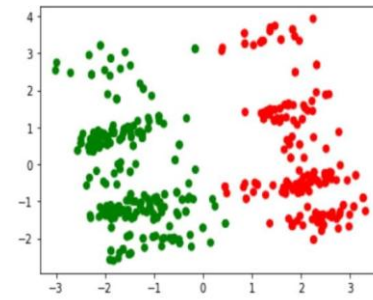


Fig 3. Scatter plotting labeled data

	q0	q1	q2	q3	q4	q5	q6	q7	q8	q9	stress	stress_per
0	1	0	1	1	5	0	0	1	0	3	0	99.157382
1	1	0	1	2	3	2	4	2	3	0	1	0.072246
2	1	2	2	2	4	0	0	2	2	3	0	99.940660
3	1	3	2	1	5	3	4	1	0	1	1	0.004008
4	1	1	2	0	4	1	0	3	2	1	0	98.599012
...	...	...	...	...	...	...	...	...	...	...	...	...
494	1	2	4	2	2	1	4	1	0	3	1	0.016305
495	0	0	2	1	4	3	4	3	1	1	1	0.009090
496	1	2	3	3	2	3	4	1	2	3	1	0.312583
497	1	3	2	2	2	4	4	1	0	3	1	0.025329
498	1	3	2	2	2	3	4	3	0	1	1	0.013639

485 rows x 12 columns

Fig 5. Stress percentage of individuals

<matplotlib.collections.PathCollection at 0x205e3cbaa60>

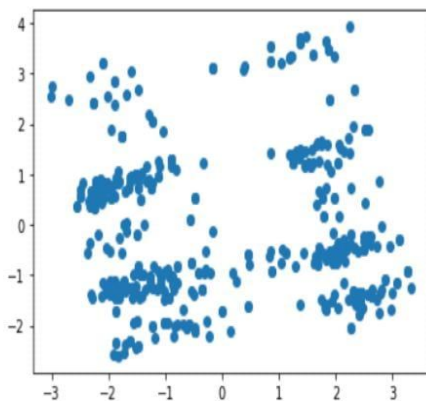


Fig 4. Scatterplot of results

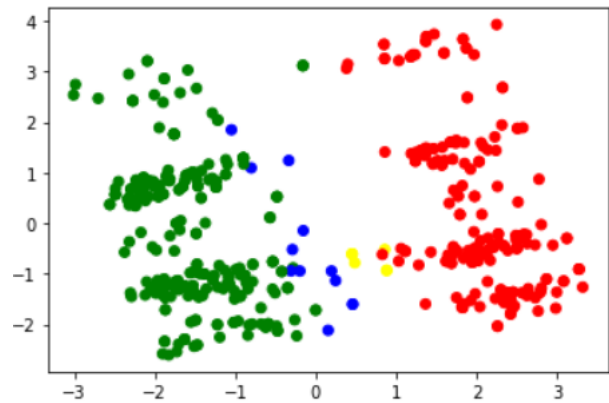


Fig 6. Stress categories of individuals

## VI. CONCLUSION

During the pandemic, most of the students were affected with higher stress levels due to many reasons like adapting to new environments of study, getting things done in a new way etc., which are the major reasons for mental stress and other psychological issues among the students. This model results in the stress percentages faced by such students. And by this we can say studying with one to one interaction is better for a student.

A dataset was constructed based on feedback from 500 students who are going through stressful situations during COVID-19. So, here some Machine learning techniques are used to compute some metrics as well as to develop clustering and classification algorithms, all of which will be helpful to enhance the mental health of future pillars of nations and make students mentally sound good.

## VII. FUTURE SCOPE

The future work is to duplicate this idea in numerous institutions for the benefit of students' mental health. A Chatbot might be integrated with the solution module.

Other Machine learning approaches may be used to acquire and analyze a considerably larger and more thorough dataset. The constraints of stress can be divided into other different sublevels.

## VIII. REFERENCES

- [1] U Srinivasulu Reddy, Aditya Vivek Thota
- [2] Norizam, Sulaiman. Determination and classification of human stress index using nonparametric analysis of EEG signals. Diss. Universiti Teknologi MARA, 2015.
- [3] Liu, David, and Mark Ulrich. "Listen to Your Heart: Stress Prediction Using Consumer Heart Rate Sensors." Online] 2014.
- [4] McBrien, J. L., Cheng, R., Jones, P. (2009).
- [5] Adnan, Nadia, et al. "University students stress level and brainwave balancing index

Copyright protected @ ENGPAPER.COM and  
AUTHORS

[Engpaper Journal](#)



<https://www.engpaper.com>