# A STATISTICAL RESEARCH ON THE EFFECTS OF MENTAL HEALTH ON STUDENTS’ CGPA

In this part we take a student mental health dataset- [data](https://www.kaggle.com/datasets/shariful07/student-mental-health)  and and try to predict the effects of student’s mental health on their CGPA. This dataset was generated by a university through google form survey to examine the current academic situation and mental health of their students. This dataset consists of various physiological and mental health attributes and more than 200 datapoints.



This task can be viewed as a tabular classification task where we can leverage the tubular techniques like decision models in order to predict the CGPA of students based on their physiological and mental health condition. We use the label encoder to convert the string labels into integer classes so that it can be fed into the models for training. Then we move on to build our model for predictions. We make use of one of the best models available i.e. XGBoost model.

## 

## CODE:

## Installation

Pip install pandas

## Imports

Import pandas as pd

from sklearn.preprocessing import OneHotEncoder, LabelEncoder

from xgboost import XGBClassifier

## Data Loading

df = pd.read\_csv('/content/Student Mental health.csv')

X= df.drop(columns=['What is your CGPA?','Timestamp'],axis=1)

y= df['What is your CGPA?']

## Data processing

X = X.apply(LabelEncoder().fit\_transform)

le = LabelEncoder()

y = le.fit\_transform(y)

## Model Building

bst = XGBClassifier()

bst.fit(X, y)

bst.score(X,y)