

# **STUDENT DEPRESSION PREDICTION**

**By K B Nandana**

# 1. Introduction

The Student Depression Prediction project is designed to assess students' mental health by analyzing various factors such as academic pressure, sleep patterns, financial stress, and dietary habits. Using machine learning, this project aims to predict the likelihood of depression among students and provide insights to help mitigate its impact.

## 2. Objectives

- Develop a predictive model to assess depression risk.
- Identify key factors contributing to student depression.
- Provide an interactive interface for data input and prediction using Streamlit.

## 3. Methodology

### 3.1

- Data Collection
- Dataset includes attributes such as academic pressure, CGPA, job satisfaction, sleep duration, dietary habits, financial stress, and family history of mental illness.
- Data preprocessing performed using NumPy and Pandas.

### 3.2

- Model Training Machine learning model trained using Scikit-Learn.
- Features engineered to optimize model performance.
- Model stored in Pickle format for deployment.

### 3.3

- Deployment
- User-friendly web application developed using Streamlit.
- Interactive UI with input fields and a prediction button.
- Aesthetic improvements using CSS styling.

## Dataset Overview

### 1. Dataset Shape

- **Rows:** 27,901
- **Columns:** 18

### 2. Column Descriptions

Column Name	Description
<b>id</b>	Unique identifier for each record
<b>Gender</b>	Student's gender (Male/Female)
<b>Age</b>	Age of the student
<b>City</b>	Location of the student

Column Name	Description
<b>Profession</b>	Occupation of the individual (mostly "Student")
<b>Academic Pressure</b>	Level of academic stress (scale of 1-5)
<b>Work Pressure</b>	Work-related stress (scale of 1-5 or 0 if not applicable)
<b>CGPA</b>	Cumulative Grade Point Average
<b>Study Satisfaction</b>	Satisfaction with academic life (scale of 1-5)
<b>Job Satisfaction</b>	Satisfaction with job (scale of 0-5, 0 if not applicable)
<b>Sleep Duration</b>	Sleep duration categories (e.g., "<5 hours", "5-6 hours", "7-8 hours")
<b>Dietary Habits</b>	Eating habits categorized as "Unhealthy", "Moderate", or "Healthy"
<b>Degree</b>	Educational qualification (e.g., B.Tech, M.Tech, BSc, etc.)
<b>Suicidal Thoughts</b>	Whether the person has had suicidal thoughts (Yes/No)
<b>Work/Study Hours</b>	Number of hours spent working or studying per day
<b>Financial Stress</b>	Financial difficulties (scale of 1-5)
<b>Family History of Mental Illness</b>	Presence of mental illness in family (Yes/No)
<b>Depression</b>	Target variable (1 = Depressed, 0 = Not Depressed)

## EDA

### 1. Univariate analysis

Academic Pressure & Work/Study Hours

- Higher academic pressure and extended study hours could be contributing factors to stress and depression.

Financial Stress

- A mean value of 3.14 suggests financial concerns may be significantly affecting students' mental health.

Depression Rate

- 58.5% of students exhibit signs of depression, indicating the need for increased mental health support and awareness programs.

#### Sleep & Diet Affecting Mental Health –

- More than 50% of students sleep less than 6 hours and 72.5% have poor dietary habits, which could contribute to stress and depression.

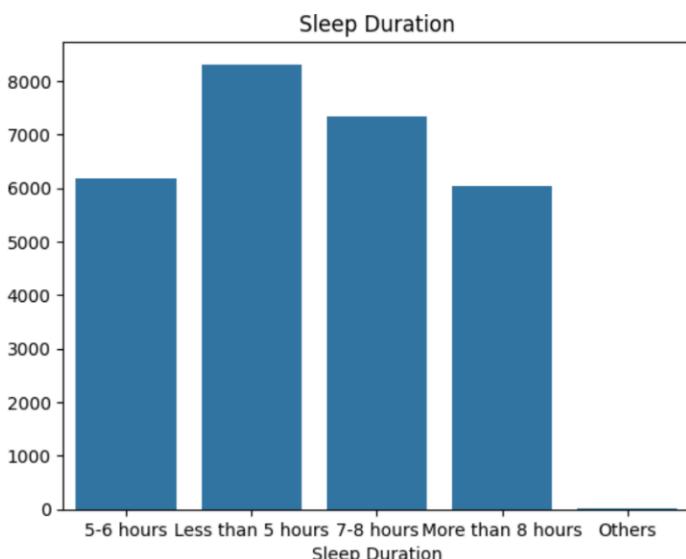
#### Academic Pressure & Financial Stress Impact

- Many students experience high academic and financial stress, which may be linked to their depressive symptoms.

## 2.Bivariate Analysis

- Academic pressure and financial stress are the biggest contributors to depression.
- Younger students are more vulnerable to depression.
- Higher study satisfaction lowers depression levels.
- Long work/study hours show a small effect on depression but are not the primary cause.

## 3.Count Plot



- The largest group in the chart consists of individuals sleeping **less than 5 hours**, which is strongly linked to an increased risk of depression.

## 4.Categorical and Numerical Analysis

### \* Gender and Depression

- Depression affects both genders almost equally.
- At least 50% of individuals in both groups report depression.

### \*Dietary and Depression

- Healthier diets are linked to lower depression rates.
- A moderate diet increases depression risk, but not as much as unhealthy diets.
- Unhealthy diets show the highest depression rates, emphasizing the importance of nutrition in mental health.
- Dietary improvements may help in depression prevention and management.

## \* Suicidal Thoughts and Depression Analysis

- Depression is significantly higher in individuals with suicidal thoughts.
  - Most people without suicidal thoughts do not experience depression.
  - The data highlights the urgent need for mental health support for those with suicidal thoughts.

```
X = df1.drop(columns=['Depression']) # Drop the target column for features
y = df1['Depression'] # Target column
✓ 0.0s Python

from sklearn.model_selection import train_test_split
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42, stratify=y)
✓ 0.1s Python

# machine learning model
# Train a Random Forest Classifier
from sklearn.ensemble import RandomForestClassifier
model = RandomForestClassifier(n_estimators=100, random_state=42)
model.fit(X_train, y_train)
y_pred = model.predict(X_test)
✓ 1.9s Python

from sklearn.metrics import accuracy_score,confusion_matrix,classification_report
print(f"Accuracy score {accuracy_score(y_test,y_pred)} \n")
print(f"Confusion matrix:{confusion_matrix(y_test,y_pred)}\n")
print(f"classification report:{classification_report(y_test,y_pred)}")
```

Class	Precision	Recall	F1-score	Support
0 (No Depression)	0.80	0.77	0.78	2313
1 (Depression)	0.84	0.86	0.85	3268
Overall Accuracy	82.31%			

- 82.31% accuracy

## Confusion matrix

- **True Positives (TP) = 2816** → Correctly predicted depression cases.
  - **True Negatives (TN) = 1778** → Correctly predicted non-depressed cases.
  - **False Positives (FP) = 535** → Incorrectly classified non-depressed as depressed.
  - **False Negatives (FN) = 452** → Missed actual depression cases

## Classification report

- **Class 0 (No Depression)**: 80% of the cases predicted as "No Depression" were actually correct.
  - **Class 1 (Depression)**: 84% of the cases predicted as "Depression" were actually correct
  - **Class 0 (No Depression)**: 77% of the actual non-depressed cases were correctly classified.
  - **Class 1 (Depression)**: 86% of actual depression cases were identified correctly.

```

features=['academic_pressure', 'cgpa', 'study_satisfaction', 'job_satisfaction',
          'sleep_duration', 'dietary_habits', 'suicidal_thoughts',
          'work_study_hours','financial_stress','family_history']

def predict_depression():
    print("Enter the following details for depression prediction:")

    academic_pressure = float(input("Enter Academic Pressure (1-5 scale): "))
    cgpa = float(input("Enter CGPA: "))
    study_satisfaction = float(input("Enter Study Satisfaction (1-5 scale): "))
    job_satisfaction = float(input("Enter Job Satisfaction (0-5 scale, 0 if not applicable): "))
    sleep_duration = int(input("Enter Sleep Duration (1 = <5 hrs, 2 = 5-6 hrs, 3 = 7-8 hrs, 4 = >8 hrs): "))
    dietary_habits = int(input("Enter Dietary Habits (1 = Unhealthy, 2 = Moderate, 3 = Healthy): "))
    suicidal_thoughts = int(input("Have you ever had suicidal thoughts? (1 = Yes, 0 = No): "))
    work_study_hours = float(input("Enter Work/Study Hours per day: "))
    financial_stress = float(input("Enter Financial Stress Level (1-5 scale): "))
    family_history = int(input("Is there a family history of mental illness? (1 = Yes, 0 = No): "))

    pred = model.predict(X_test)[0]

    if pred == 1:
        print("Prediction: The person is likely to have depression.")
    else:
        print("Prediction: The person is unlikely to have depression.")

```

## 4. Implementation Details

### 4.1

- Streamlit UI Features Attractive UI: Gradient background, responsive columns, and custom fonts.
- User Input Fields: Interactive number inputs for various depression indicators.
- Prediction Output: Displays a prediction of depression likelihood based on user inputs.

### 4.2

- CSS Enhancements Custom gradient headers with icons.
- Improved column layout for better organization.
- Dark-themed prediction box with pink text for better readability..

## 5. Conclusion

This project successfully demonstrates a data-driven approach to identifying student depression risk. By leveraging machine learning and interactive UI, the system offers valuable insights into student mental health.