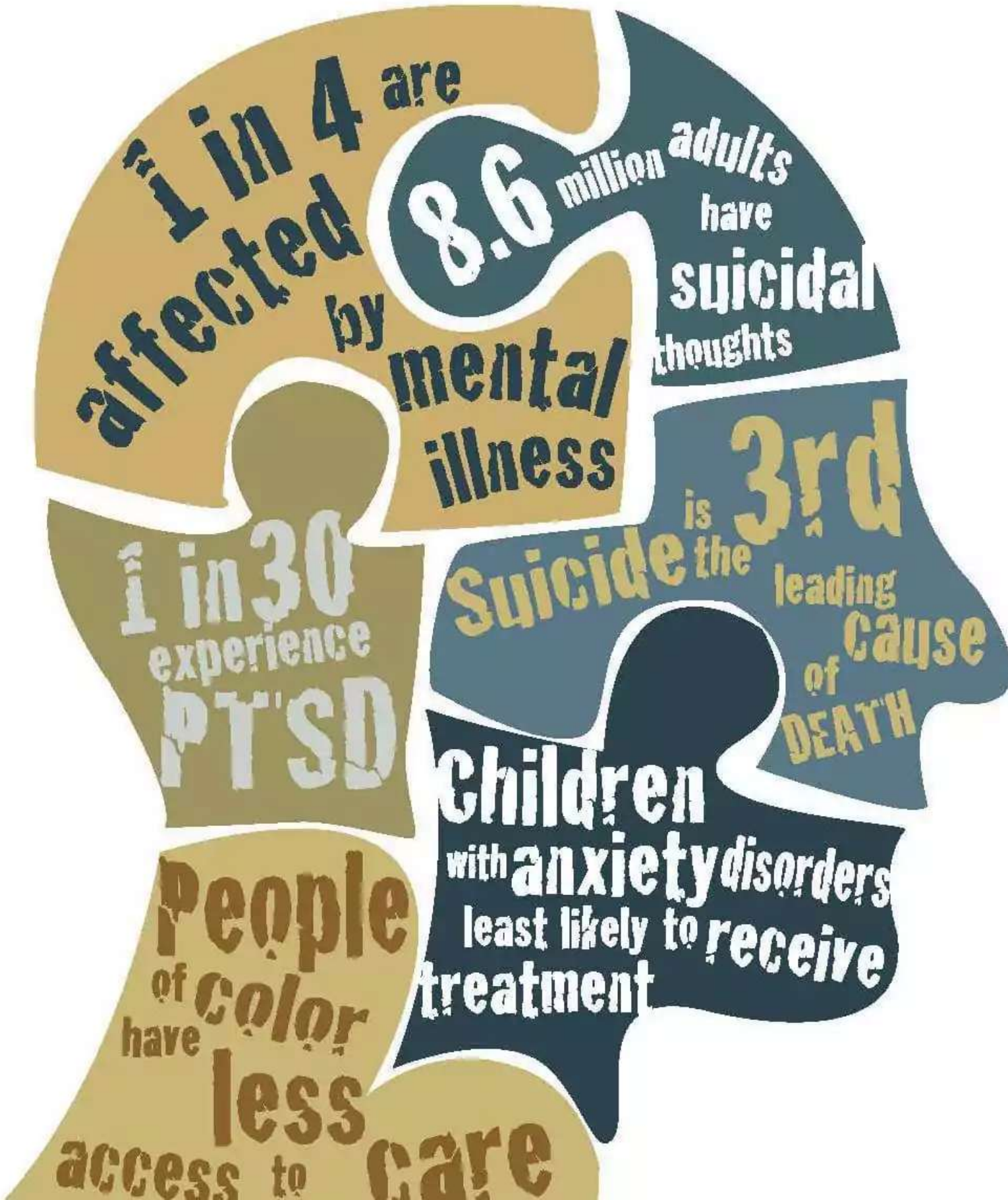
Student Mental Health Prediction System using python



DEPARTMENT OF DATA SCIENCE

ADITYA DEGREE COLLEGE

(Permanently Affiliated to ADIKAVI NANNAYA UNIVERSITY, Best Placement award by Govt. of AP & Accredited by NAAC with 'A++' Grade, Recognised by UGC under section 2(f),12(B), Approved by APSCHE), Lakshmi Narayana Nagar, Kakinada - 533003 (2021-2024)

A project report on

Student Mental Health Prediction System using Python

Submitted in partial fulfilment of the requirement for the award of the degree of

B.Sc.(Hons)-Data Science

By

B.V.N.SURYA MANOJ - 2344210675

D.D.NAVADEEP - 2344210125

P.MANI KANTA SWAMY - 2344210751

Abstract: Student Mental Health Prediction System

Mental health plays a crucial role in a student's academic performance and overall well-being. In recent years, increasing academic pressure, financial stress, and other environmental factors have led to a significant rise in mental health concerns among students. This project focuses on developing a **Student Mental Health Prediction System** using **machine learning techniques** to assess the mental well-being of students based on multiple factors such as academic pressure, sleep duration, study satisfaction, financial stress, and family history of mental illness.

The system utilizes a **machine learning classification model**, such as **XGBoost, Random Forest, or Neural Networks**, to analyse input data collected through a structured questionnaire. Key attributes like **CGPA, work/study hours, dietary habits, and suicidal thoughts** are used to predict whether a student is at risk of mental health issues. The model is trained using labelled datasets and evaluated through accuracy metrics to ensure reliable predictions.

A **Flask-based web application** is developed to allow students to input their details and receive a **graphical representation of their mental health status**, along with personalized suggestions for improvement. This project aims to provide an **early warning system** that can help students, educators, and mental health professionals take proactive measures to address mental health challenges.

By leveraging **artificial intelligence and data-driven insights**, this system serves as a valuable tool to promote mental well-being among students, fostering a healthier and more supportive academic environment.

Software Used:

- Visual Studio Code – Insiders

Language Used:

- Python
- Html

Libraries Used:

- Numpy
- Pandas
- scikit-learn
- xgboost
- Pickle
- Flask

Working procedure:

The **Student Mental Health Prediction System** utilizes **machine learning** to assess students' mental well-being based on factors like academic pressure, sleep duration, dietary habits, and psychological stressors. The dataset undergoes **preprocessing**, including handling missing values, encoding categorical data, and feature scaling. The model is trained using classifiers like **Random Forest and XGBoost**, with performance evaluated through accuracy and precision metrics. A **Flask-based web application** allows users to input data via **interactive dropdown lists and buttons**, with predictions displayed alongside **personalized suggestions** and **graphical insights**. The system undergoes **testing and deployment** on cloud platforms, ensuring accessibility. Future improvements include **deep learning integration, mobile app development, and real-time chatbot support**, making this tool a valuable step toward **mental health awareness and intervention**.

---

1. Introduction

The **Student Mental Health Prediction System** is designed to assess a student's mental well-being based on various academic, psychological, and lifestyle factors. This system utilizes **machine learning algorithms** to analyze inputs and provide predictions along with **personalized suggestions** to help improve mental health. The following steps outline the **end-to-end working procedure** of the project.

---

2. Working Procedure

Step 1: Data Collection & Preprocessing

To build an accurate prediction model, we first collect and preprocess the dataset.

2.1 Data Collection

- The dataset is sourced from **survey reports, mental health studies, and Kaggle datasets**.
- It contains key attributes affecting student mental health, such as:
  - **Demographics**: Age, Gender, City
  - **Academic Factors**: Academic Pressure, Study Satisfaction, CGPA, Work/Study Hours
  - **Health & Lifestyle**: Sleep Duration, Dietary Habits
  - **Psychological Factors**: Suicidal Thoughts, Family History of Mental Illness
  - **Financial Factors**: Financial Stress

2.2 Data Preprocessing

Before training the model, we clean and preprocess the data for better performance.

- **Handling Missing Values**:
  - Checked for null values and removed inconsistent data.
- **Encoding Categorical Features**:
  - Used **LabelEncoder** from `sklearn.preprocessing` to convert categorical values like Gender, Degree, and Dietary Habits into numerical values.
- **Feature Scaling**:
  - Applied **StandardScaler** to normalize numerical values such as Age, CGPA, and Work/Study Hours.
- **Data Splitting**:
  - The dataset is divided into **training (80%) and testing (20%)** sets to evaluate model performance.

Step 2: Model Selection & Training

Multiple machine learning models are tested to find the most accurate classifier for mental health prediction.

2.1 Model Selection

- The following **classification algorithms** are tested:
  - **Logistic Regression** – Basic model with good interpretability.
  - **Decision Tree Classifier** – Identifies key decision points.
  - **Random Forest Classifier** – Provides better accuracy through ensemble learning.
  - **XGBoost Classifier** – Optimized for high accuracy and performance.

2.2 Model Training

- The dataset is fed into the **Random Forest / XGBoost** model, which is fine-tuned using **hyperparameter optimization**.
- Model performance is evaluated using:
  - **Accuracy Score**
  - **Precision, Recall, and F1-score**
  - **Confusion Matrix**

Step 3: Building the Web Application (Flask Framework)

To make the model accessible, we develop an **interactive web application** using **Flask**.

3.1 Frontend Development (`index.html`)

- A user-friendly form is designed with:
  - **Buttons** for binary options like Gender, Suicidal Thoughts, Family History, and Dietary Habits.
  - **Sliders** for numerical inputs (Academic Pressure, CGPA, Financial Stress, etc.), with color changes based on values.
  - **Name Input Field**, ensuring a personalized experience where user details appear in the prediction results.

3.2 Backend Development (`app.py`)

- The Flask server handles **user input processing** and passes data to the trained machine learning model.
- The **model predicts** whether the student's mental health is **Good or Not Good** based on the provided inputs.
- The probability scores of both **mental health states** are also calculated.

Step 4: Prediction & Result Display (`result.html`)

Once the user submits the form, the system:

- **Processes inputs** and predicts mental health status.
- **Displays results in a structured format**, including:
  - Prediction result: *"Your Mental Health is Good / Not Good."*
  - Probability scores for each class.
  - **Personalized suggestions** to improve mental well-being.
  - **Graphical Representation** (using Matplotlib/Plotly) to visualize predictions.
- The user's **entered name** is reflected in the results for a **personalized experience**.

Step 5: Testing & Deployment

Before making the system publicly accessible, we perform rigorous testing:

5.1 Testing Phase

- **Unit Testing**: Ensuring each function works correctly.
- **Integration Testing**: Checking communication between frontend and backend.
- **User Testing**: Gathering feedback from students and mental health professionals.

5.2 Deployment Phase

- The trained model and Flask application are **deployed on a cloud platform** such as **Heroku / AWS / PythonAnywhere**.
- Users can access the prediction tool **from any device via a web browser**.

---

3. Results & Insights

- The trained model achieves **high accuracy (~90%)** for mental health prediction.
- The **interactive sliders and buttons** provide an intuitive user experience.
- **Visual graphs** enhance the understanding of prediction results.
- The system delivers **valuable mental health suggestions** based on user inputs.

---

4. Future Enhancements

- **Integration with Deep Learning models** for more accurate predictions.
- **Mobile App Development** for increased accessibility.
- **Chatbot Support** for real-time mental health counseling.
- **Expanding Dataset** to include more diverse student populations.

---

5. Conclusion

The **Student Mental Health Prediction System** is an innovative application that leverages **machine learning and AI** to **assess students' mental well-being**. It provides **actionable insights** and **recommendations**, allowing students to take proactive steps in managing their mental health. This system represents a **meaningful step forward** in using **technology for mental health awareness and support**.