Microsoft

Mental Health Monitor

Sudhanshu Tiwari

Anshul Kumar

Avisha Goyal

Janeika J

Microsoft Learn

STUDENT AMBASSADOR

Agenda

PROBLEM STATEMENT

Mental Health Monitoring: Develop an AI system that monitors mental health conditions by analyzing speech patterns, facial expressions, and text-based data. The system can provide insights and alerts to both patients and healthcare professionals, facilitating early intervention and personalized treatment.

SOLUTION

The proposed solution is an AI-enabled Mental Health Monitoring system that analyzes speech patterns, facial expressions, and text-based data to detect early signs of mental health conditions. The system provides real-time insights and alerts to both patients and healthcare professionals, facilitating personalized treatment and proactive intervention. The ultimate goal is to **empower individuals, reduce stigma, and improve the overall quality of life** for those facing mental health challenges

Introduction

As college life brings a unique blend of challenges and opportunities, it's crucial to address the mental health needs of students. Our project emerges at the intersection of technology and student well-being, offering a dynamic approach that caters to the specific demands of this phase.

More college students are seeking therapy or counseling as anxiety, depression, and suicidal thoughts reach record levels, according to a recent report (2021-22)

The study, based on surveys from 96,000 students across 133 campuses, revealed that 46% experienced depression symptoms, 39% had anxiety disorders, and 15% seriously contemplated suicide in the past year, marking the highest rates in the survey's 15-year history.

Introduction

Youth are more likely to suffer from depression due to :
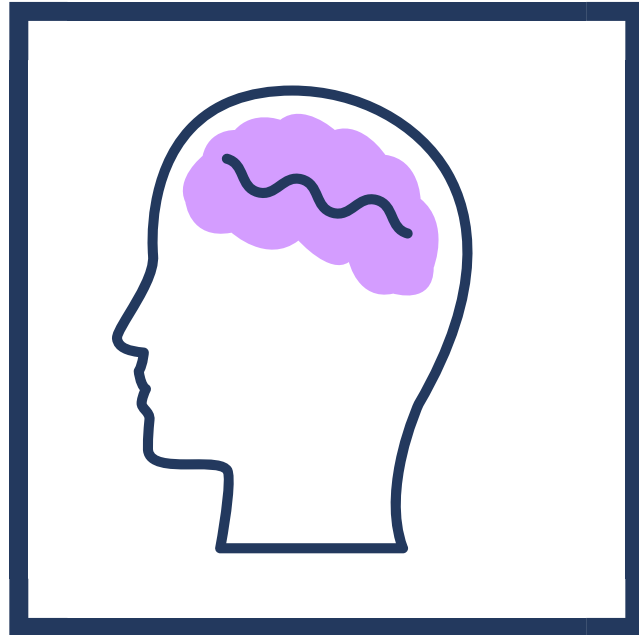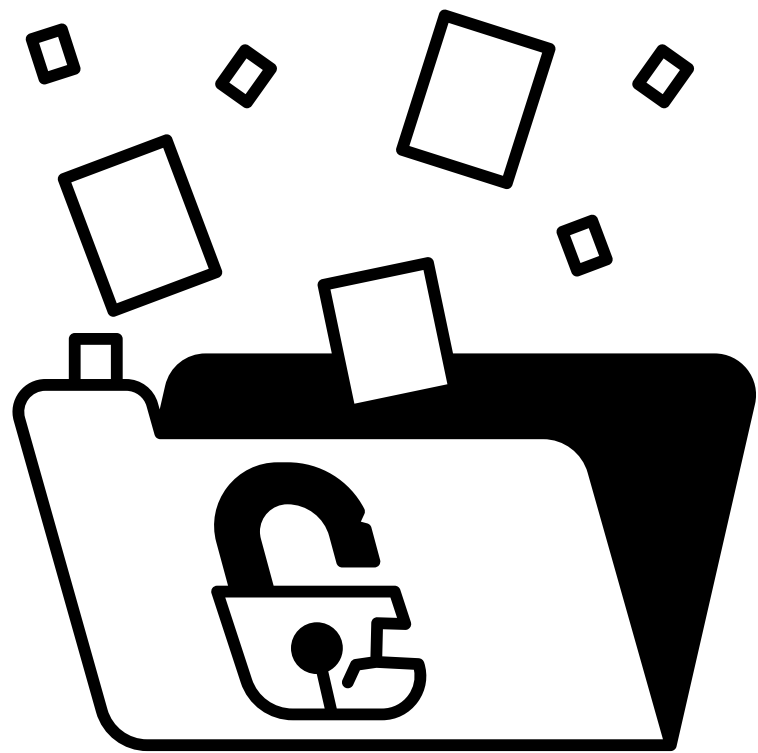

1. **Imbalances in brain chemicals during puberty, such as serotonin, norepinephrine, and dopamine, can lead to depression.**

2. **The stress of societal expectations and academic pressure can contribute to the development of depression.**

MODEL SELECTION PROCESS

1.Data Collection

- The dataset used in the project was scraped from Reddit and comprises 232,074 rows, with an equal distribution between two classes - "suicide" and "non-suicide." This balanced class distribution, where each class has an equal number of samples, is beneficial for training machine learning models as it helps prevent bias and ensures that the model learns from both classes equally.

- It aims to develop a system for classifying posts as suicide-related or not, helping identify and support individuals discussing mental health issues on social media.

2.Text Preprocessing

- Text processing is performed to create a cleaned dataset for the Suicide and Depression Detection project. The raw dataset, containing text data, is loaded and undergoes various preprocessing steps, such as lowercasing, tokenization, removing punctuation, eliminating stop words, lemmatization or stemming, handling numbers, and removing irrelevant elements like HTML tags, URLs, and emojis.

- By executing it, the text data is properly processed, cleaned, and prepared for analysis and model training. The resulting cleaned dataset forms a reliable foundation for building an accurate and robust system to detect signs of suicidal or depressive tendencies in text inputs.

3. Data Cleaning

- In this suicide detection project, the data underwent preprocessing to convert text to lowercase, tokenize, remove punctuation and stop words, and encode labels. Data cleaning involved handling missing data, inconsistencies, duplicates, and outliers.

- The dataset was balanced, with approximately 4:6 class distribution for suicide and non-suicide instances. The final dataset had 174,436 rows and was split into training, testing, and validation sets with an 8:1:1 ratio. This clean and prepared dataset will be used for analysis and model training.

4.Exploratory Data Analysis (EDA)

- This analysis helps identify patterns, trends, and potential issues, guiding the necessary text preprocessing and data cleaning steps. It aids in detecting class imbalances, missing data, and other anomalies that could impact model performance.

- This phase provides deeper insights into the dataset's transformed state, including the distribution of word frequencies, the effectiveness of feature engineering, and the impact of class balancing techniques. These insights are critical in building better feature representations and designing more robust models for accurate suicide detection.

5. Representation Learning

- In the project, we experimented with pre-trained Twitter GloVe embeddings to enhance representation learning. These embeddings are word vectors capturing semantic relationships and word meanings from a large Twitter corpus. We downloaded and stored them for use in the model. By leveraging pre-trained embeddings, the model gains a better understanding of text, improving its performance in suicide detection by detecting meaningful patterns and relationships.

6. Model Building & Evaluation

Model building involves selecting an algorithm, optimizing parameters, and refining features for effective machine learning. Evaluation, crucial for assessing generalization, employs metrics like accuracy and precision. Techniques such as k-fold cross-validation ensure robustness. Balancing model complexity prevents overfitting. Continuous refinement, guided by thorough evaluation, is essential for creating reliable and accurate predictive models.
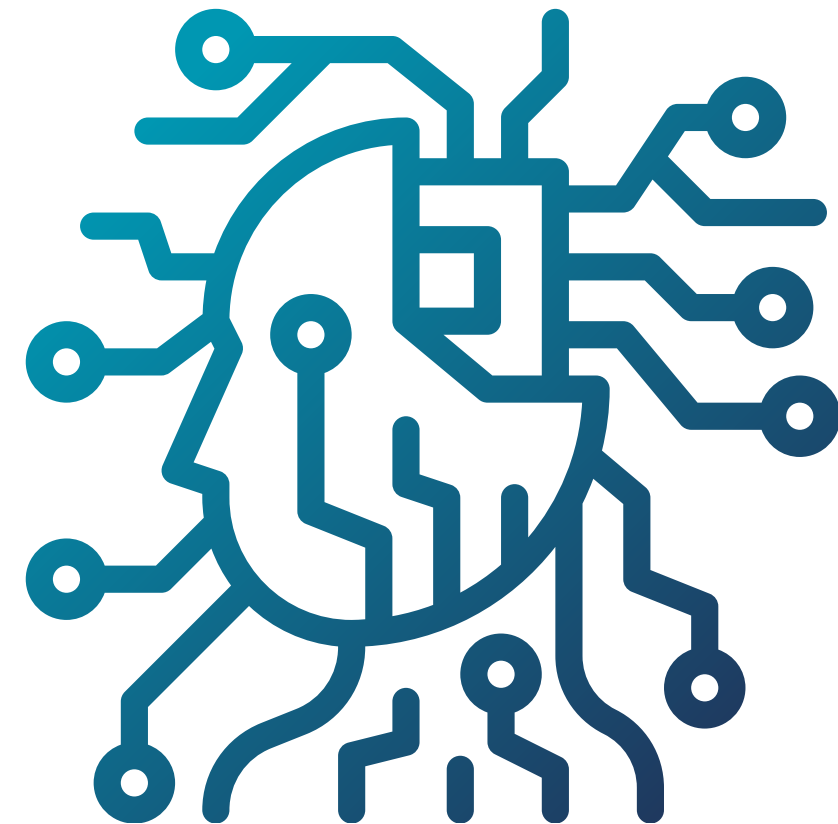
7. Model Selection

Model selection is a critical aspect of machine learning, involving the careful choice of algorithms and architectures to address specific tasks. It requires assessing various models based on factors like performance, computational efficiency, and interpretability. Techniques such as cross-validation help determine the model's ability to generalize to new data. Striking the right balance between simplicity and complexity is key in choosing a model that best suits the problem at hand, ensuring optimal performance and meaningful insights.
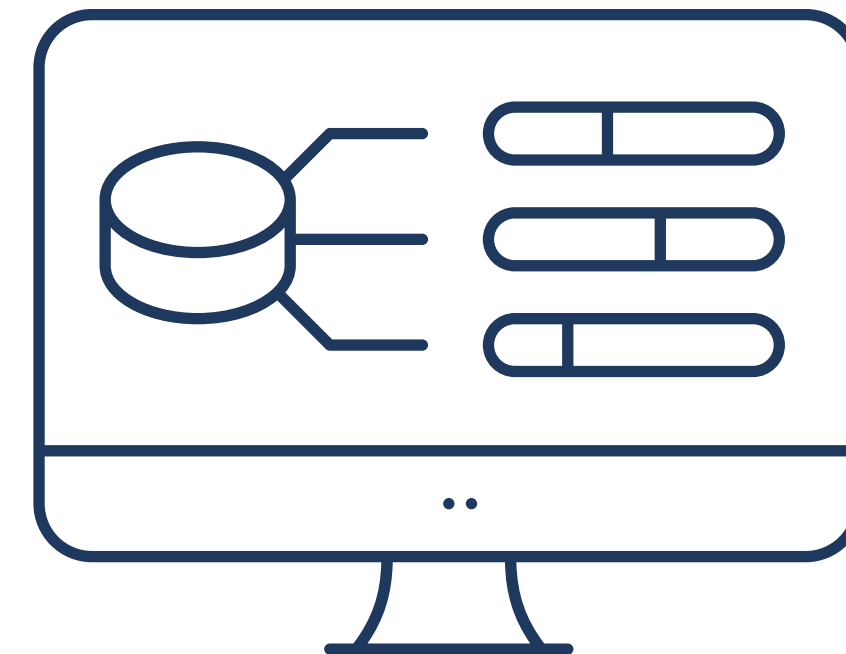
8. Chatbot Integration

- The project integrated a chatbot with the suicide detection model. Using DialoGPT, the chatbot responds naturally in conversations. Custom retrieval-based responses were included to address mental health concerns. When a user's input indicates distress, the suicide detection model is triggered to assess their risk level.

- The chatbot can then provide appropriate support, such as helpline numbers or encouraging professional help. The integration offers empathetic assistance, detecting signs of distress and guiding users to seek necessary support.