



## GROUP 29

# Analyzing Social Media Posts for Mental Health Disorder Detection

ROLL NUMBER	NAME
13000121033	SOUMYADEEP NANDY
13000121037	PRITHWISH SARKAR
13000121040	SAGNIK MUKHOPADHYAY
13000121058	ARKAPRATIM GHOSH

PROJECT-II ( PROJ-CS781 )

CSE : SEMESTER 7

# CONTENT

1. Motivation
2. Introduction
3. Research Work
4. Problem Definition
5. Proposed Workflow
6. Implementation
7. Results and Analysis
8. Prototype
9. Conclusion
10. What's Next ?
11. References

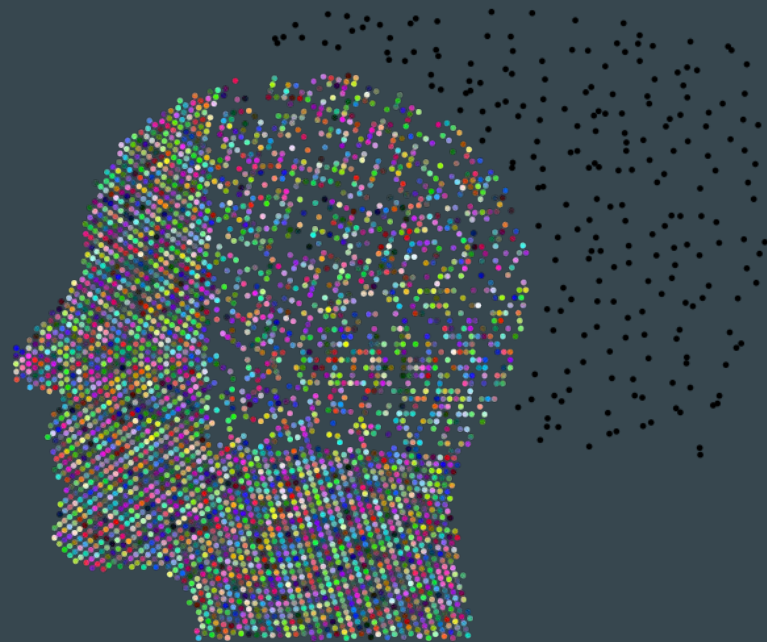


# MOTIVATION

- ***Rising global concern over mental health disorders*** : Mental health issues are affecting millions worldwide, requiring urgent attention .
- ***Social media is a key outlet for emotional expression*** : Platforms like *Twitter* and *Reddit* reveal mental health struggles in real-time.
- ***Early detection of mental health issues can save lives*** : Identifying mental health disorders early helps provide timely interventions.
- ***Machine learning can automate detection of mental health disorders*** : Technology enables efficient analysis of large social media data for early warning signs.
- ***Potential to assist mental health professionals and organizations*** : Provides valuable insights for mental health monitoring and public health efforts.
- ***Opportunity to improve mental health awareness on social platforms*** : Can support campaigns that foster awareness and reduce stigma online.
- ***Lack of a publicly available application*** which incorporates text, image, video, audio, emotions and situations derived from images/frames for overall mental issue classification and corresponding wellbeing mapping.

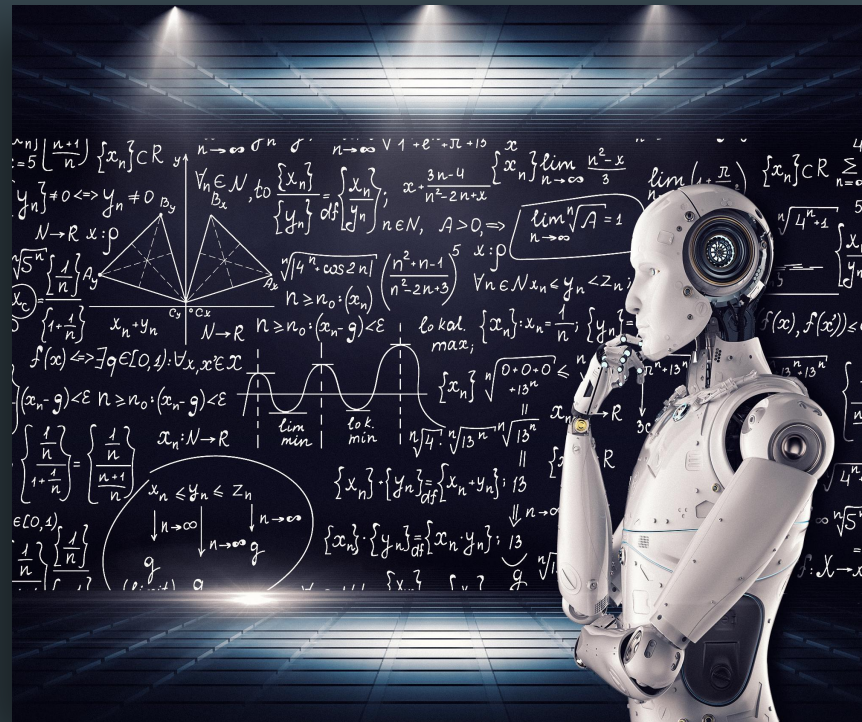
# INTRODUCTION

- **Role of social media in mental health expression** : People share emotions, struggles, and experiences on platforms like Twitter and Reddit.
- **Goal of the project** : To detect mental health disorders early through the analysis of social media posts or by uploading images, videos, facial expression recognition, generating image captions ,manually inserting text and retrain model in the process.
- **Leveraging machine learning** : Using advanced techniques like NLP and classification models to analyze text normally posted, extracted from image, audio and video.



# INTRODUCTION (CONTINUED)

- ➔ **Focus on text classification for the base model :** Analyzing language patterns to classify posts related to mental health issues.
- ➔ **Impact of early detection :** Can enable timely intervention and direct users to mental health support services.
- ➔ **Models used in the project :** Techniques like Logistic Regression, XGBoost are applied for high accuracy. Ensemble Model using the individual models is created to get higher accuracy and generalization.



# INTRODUCTION (CONTINUED)

- ➔ **Broader Goal:** Use technology to assist mental health professionals and enhance public health awareness.
- ❑ **Implement text analysis** to detect mental health concerns from user inputs.
- ❑ **Develop an image analysis** system for text extraction, facial expression recognition, and gesture analysis.
- ❑ Create a **video analysis module** for converting video to audio, extracting text, and recognizing expressions.
- ❑ Incorporate **image captioning** using transformers to contribute to mental wellbeing insights.
- ❑ Build a system to **fetch and analyze Reddit and Twitter user posts** for mental health classification.
- ❑ Enable **multilingual analysis** by translating non-English content to English.
- ❑ Use generative AI to **map mental health classifications to Ryff's Scale** of Psychological Well Being.
- ❑ Provide **data visualization** for emotional trends and post analysis.
- ❑ Develop an **ensemble model with multiple algorithms** for robust mental health classification.
- ❑ Integrate a feature to **retrain the model with user-provided data** for continuous improvement.

# RESEARCH WORK

## → *Social media and mental health research*

Explored studies on how social media data can reveal mental health conditions.

## → *Key study by Choudhury et al. (2013)*

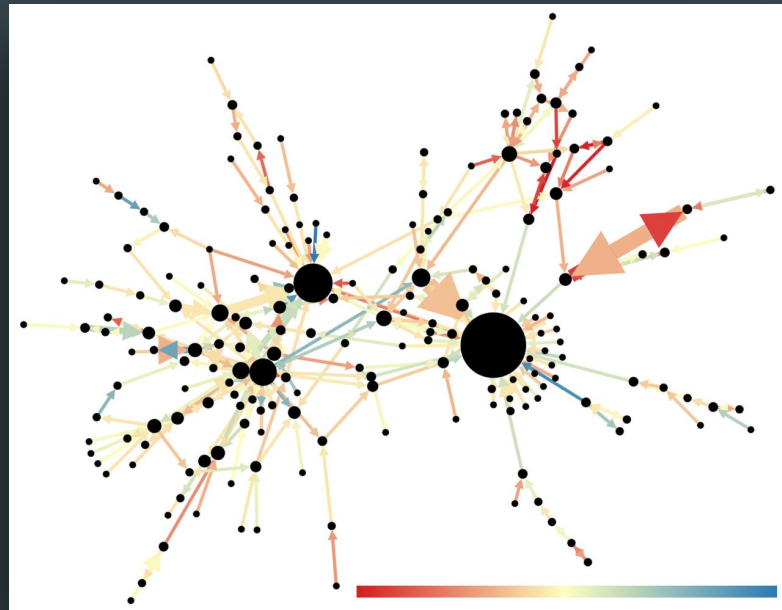
Showed the predictive power of Twitter data in identifying depression through linguistic patterns.

## → *Guntuku et al. (2017) review*

Synthesized various approaches to detecting mental illness using sentiment analysis on social platforms.

## → *Mathur et al. (2022) systematic review*

Highlighted the success of machine learning techniques in detecting disorders like depression and anxiety.





# RESEARCH WORK (CONTINUED)

## → *Nadeem (2016) study on Twitter*

Demonstrated the potential of text analysis to identify at-risk individuals based on emotional cues in tweets.

## → *Al Sagri and Ykhlef (2020) approach*

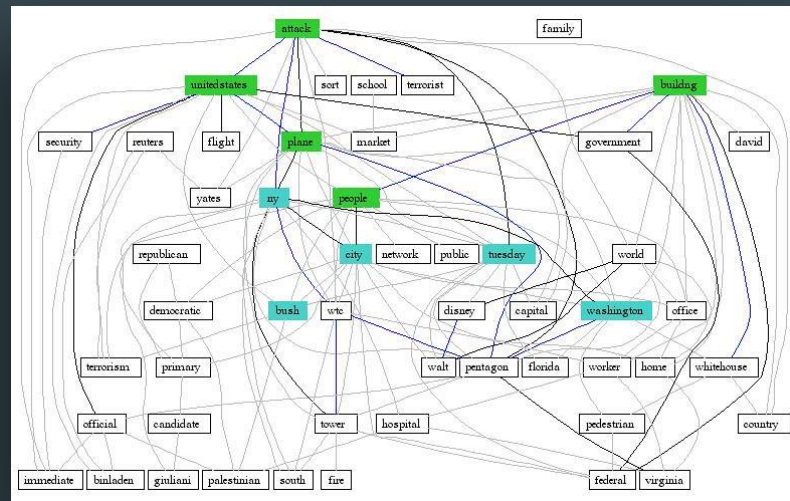
Combined linguistic and behavioral features for more accurate depression detection on Twitter.

## → *Recent study by Vaishnavi et al. (2022)*

Comparative analysis of algorithms to identify mental health conditions from social media posts.

## → *Ethical considerations by Safa et al. (2023)*

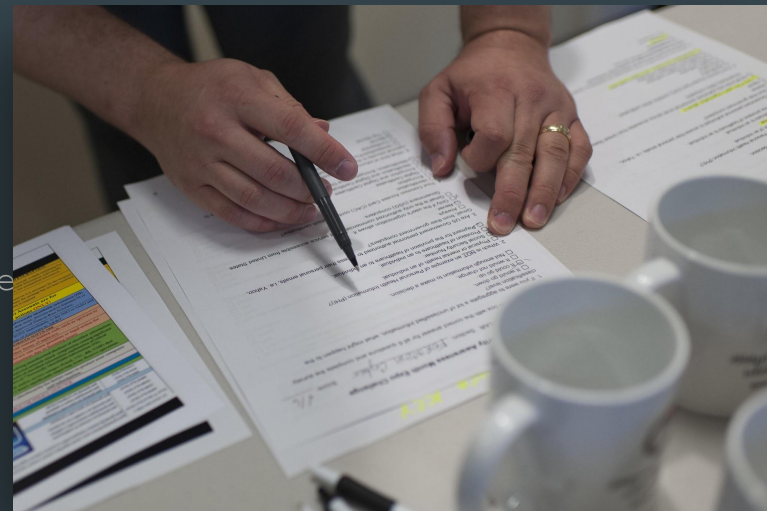
Addressed data privacy challenges in mental health detection research using social media data.





# PROBLEM DEFINITION

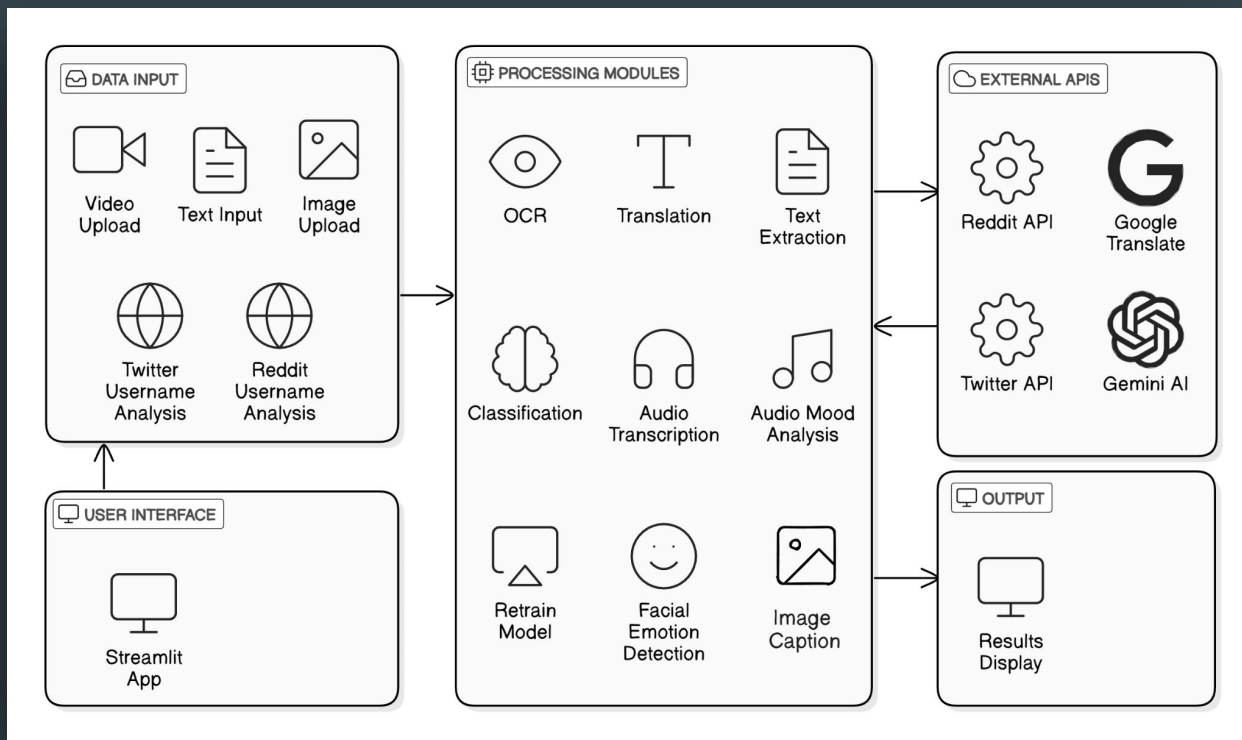
- **Rising prevalence of mental health disorders** : Increasing cases of depression, anxiety, PTSD, and other mental health issues globally.
- **Challenges in early detection** : Mental health problems are often undiagnosed until advanced stages, limiting timely intervention.
- **Vast amount of unstructured social media data** : Social media platforms generate large volumes of text that can indicate mental health struggles.
- **Need for efficient detection methods** : Manual analysis of social media posts is time-consuming; automation using machine learning is essential.
- **Goal** : To develop a system that accurately classifies social media posts based on mental health disorders.



# PROPOSED WORKFLOW

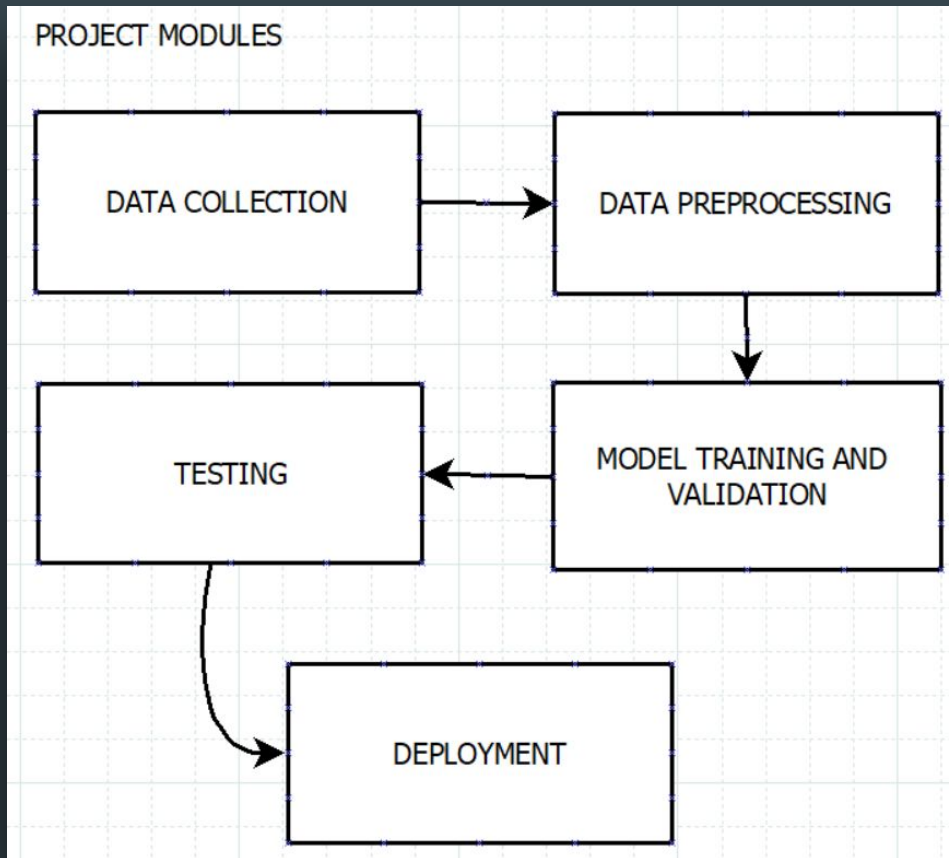
- ➔ **Data Collection** : Collect Reddit posts using the PRAW API and extract labeled data for mental health categories like Normal, Anxiety, Depression, PTSD, and Bipolar.
- ➔ **Data Preprocessing** : Clean the collected data by removing URLs, stopwords, and special characters. Normalize and tokenize the text, followed by converting it into numerical formats ( *TF-IDF*, *Bag Of Words* ) for analysis.
- ➔ **Model Training and Evaluation** : Train machine learning models including Logistic Regression, Naive Bayes, SVM, Random Forest, XGBoost, KNN, and LSTM. Evaluate their performance using metrics such as accuracy, precision, recall, and F1-score. Apply Hyperparameter Tuning on ML models as needed to improve the accuracy further. Leverage Ensemble Learning with two models to get higher accuracy
- ➔ **Testing and Deployment** : Test the best-performing models and deploy them on a user-friendly interface using Streamlit. Ensure the system supports real-time classification for various inputs like text, image, video and user profiles from social media platforms like Reddit and Twitter.

# PROPOSED WORKFLOW (CONTINUED)



SYSTEM OVERVIEW

# PROPOSED WORKFLOW (CONTINUED)



# IMPLEMENTATION

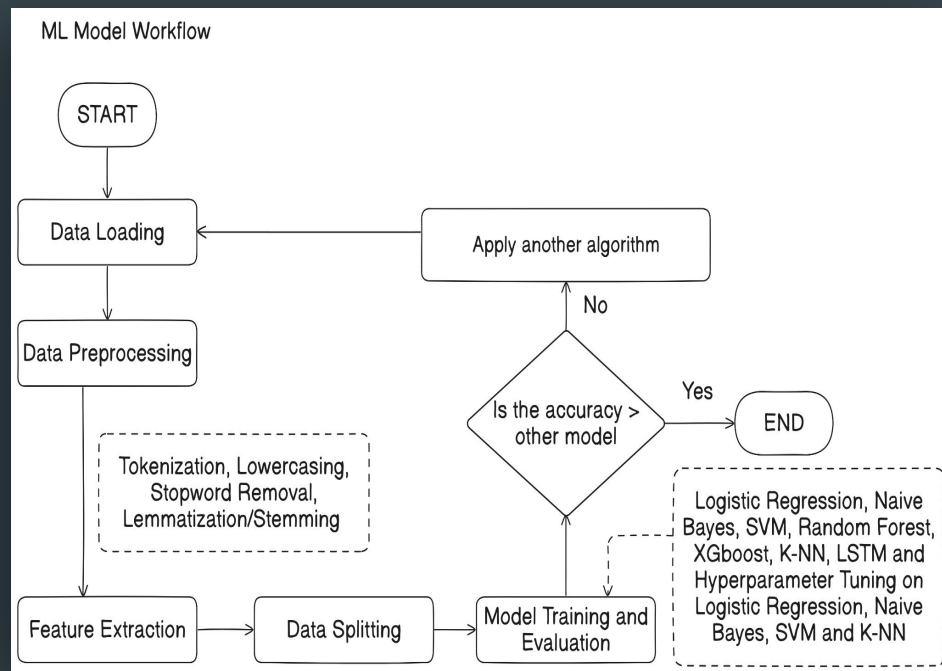
- **Data Collection** : Scraped Reddit posts related to mental health issues and for analysis.
- **Data Preprocessing** : Cleaned and normalized the text by removing URLs, stop-words, punctuation, and applied tokenization and lemmatization techniques.
- **Feature Extraction** : Utilized Bag of Words, Term Frequency-Inverse Document Frequency (TF-IDF) to convert text into numerical format for machine learning models.
- **Splitting the Dataset** : Divided the dataset into training and testing sets to train models and evaluate their performance.

Rqmt ID	Requirement Item	Requirement Analysis Status
FR-001	Collect social media data from Reddit.	Completed ▼
FR-002	Implement data cleaning and preprocessing.	Completed ▼
FR-003	Train machine learning and deep learning models.	Completed ▼
FR-004	Evaluate models using performance metrics (accuracy, recall, F1 Score, Support).	Completed ▼
NFR-001	Testing different features of the web application	Completed ▼
NFR-002	Final Web Application Deployment	Completed ▼

## FUNCTIONAL REQUIREMENTS

# IMPLEMENTATIONS (CONTINUED)

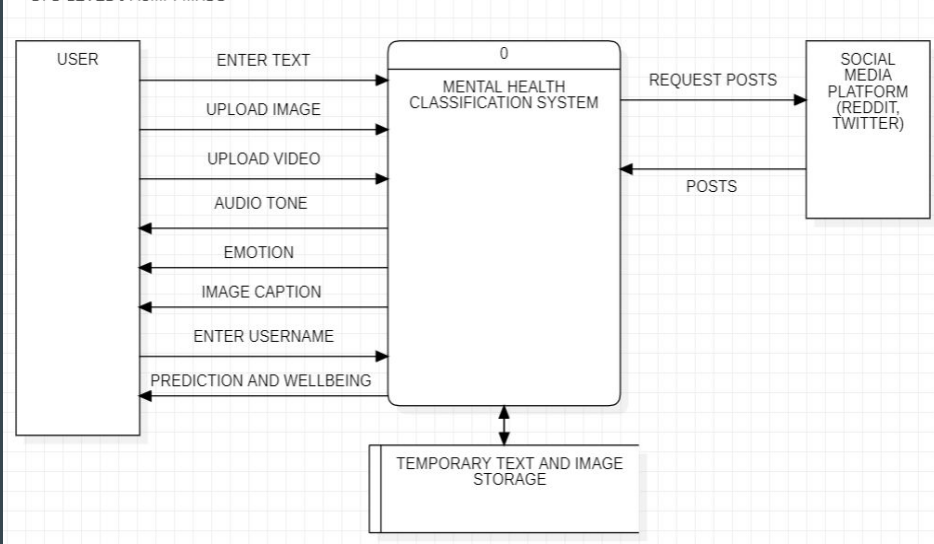
- **Model Training** : Trained multiple models: Logistic Regression, k-Nearest Neighbors (k-NN), Support Vector Machine (SVM), Naive Bayes, Random Forest, XGBoost, Long Short Term Memory and Ensemble Model
- **Hyperparameter Tuning** : Applied RandomizedSearchCV to optimize the performance of each model.
- **Model Evaluation** : Used metrics like accuracy, precision, recall, F1-score, and confusion matrices to evaluate model effectiveness.
- **Prediction and Deployment** : Implemented the best-performing model for predicting mental health issues from social media posts and deployed using Google Colab.



MODEL WORKFLOW

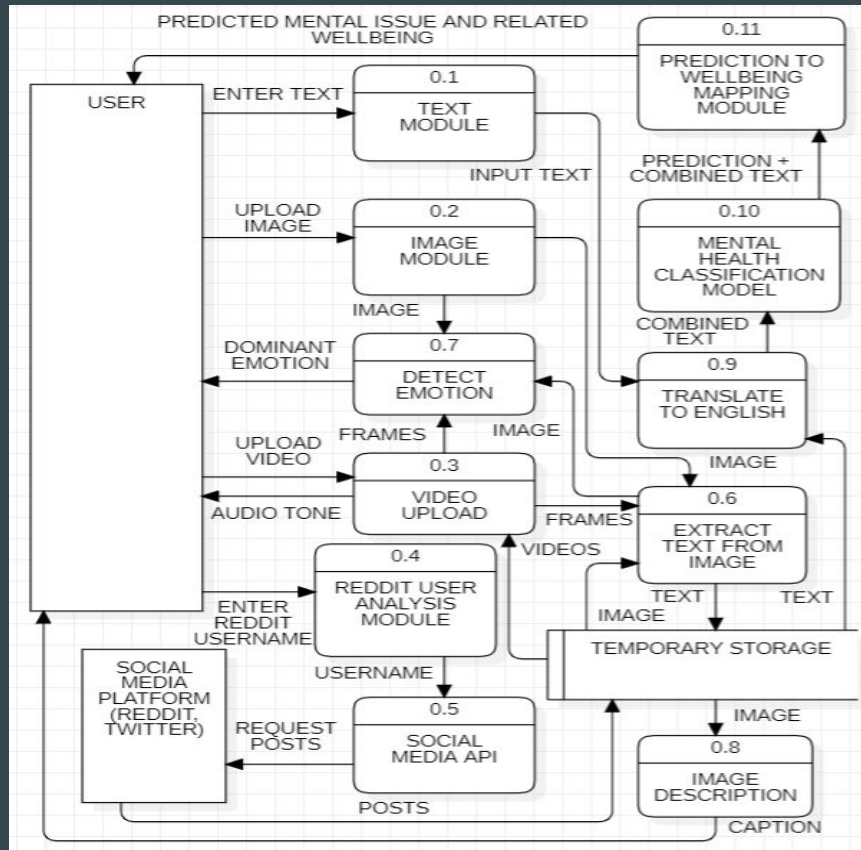
# IMPLEMENTATION (CONTINUED)

DFD LEVEL 0 ASMPFMHDD



DFD LEVEL 0

DFD LEVEL 1

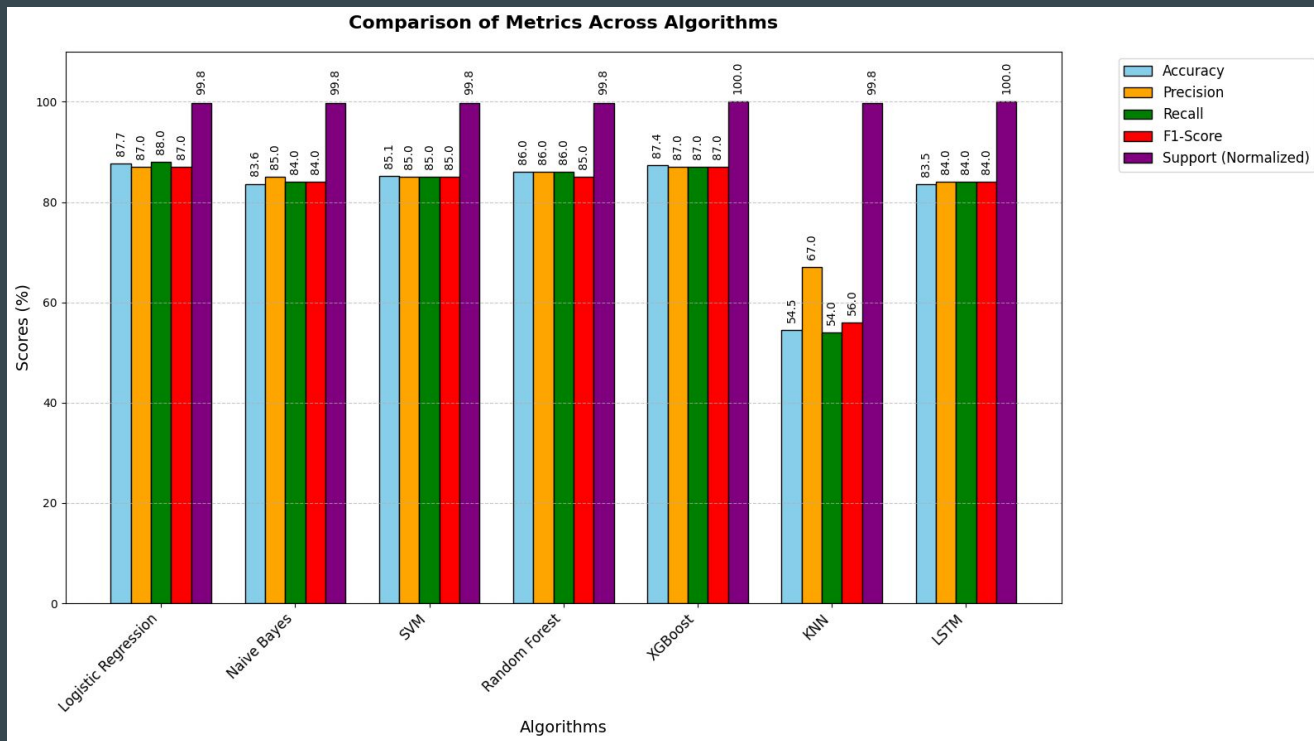




# RESULTS AND ANALYSIS

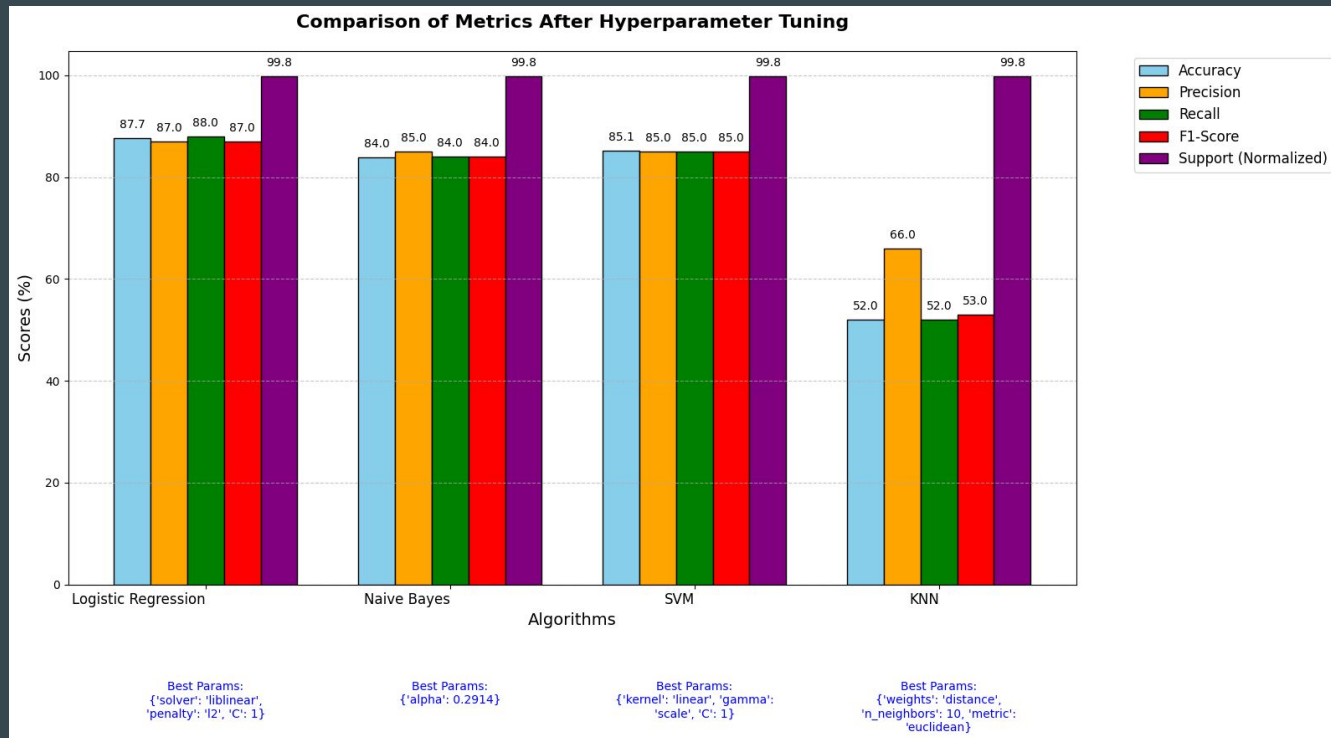
- **Logistic Regression** : Achieved consistent performance with highest accuracy. Precision and recall indicate reliable classification for balanced datasets.
- **Naive Bayes** : Performed well with text data, especially for independent features. Precision was slightly lower for imbalanced classes but remained effective overall.
- **Support Vector Machine (SVM)** : Delivered high accuracy for nonlinear classification tasks. The model showed robustness with complex feature interactions.
- **Random Forest** : Provided strong classification performance with reduced overfitting. Achieved balanced precision and recall across all mental health classes.
- **XGBoost** : Delivered the highest accuracy and efficiency. The model demonstrated excellent handling of imbalanced datasets with robust predictions.
- **LSTM** : Outperformed traditional methods for sequential data. Captured contextual and temporal information effectively, achieving competitive results with complex text patterns.
- **K-nearest Neighbours** : Performed poorly with the lowest accuracy among all.

# RESULTS AND ANALYSIS (CONTINUED)



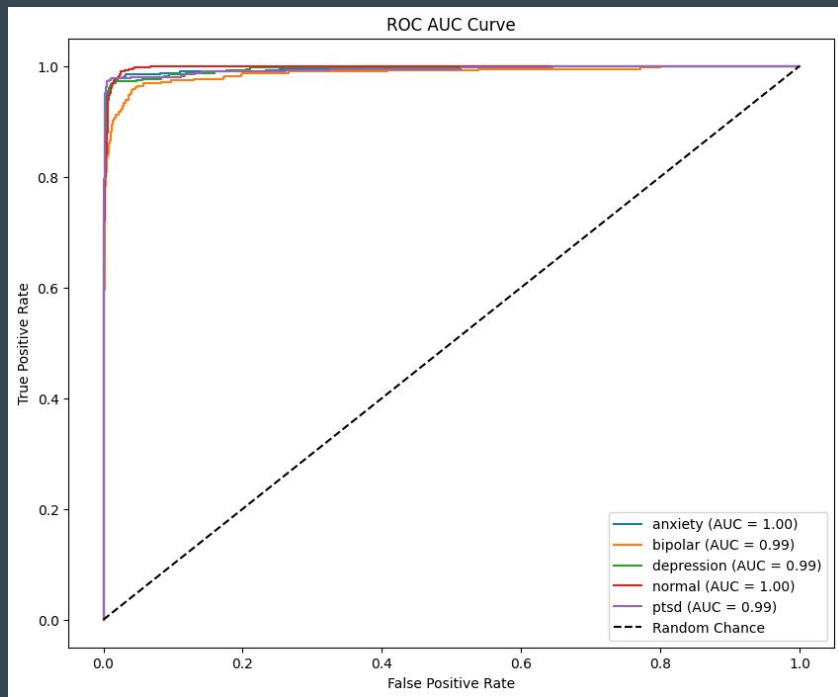
RESULT COMPARISON OF DIFFERENT MODELS

# RESULTS AND ANALYSIS (CONTINUED)

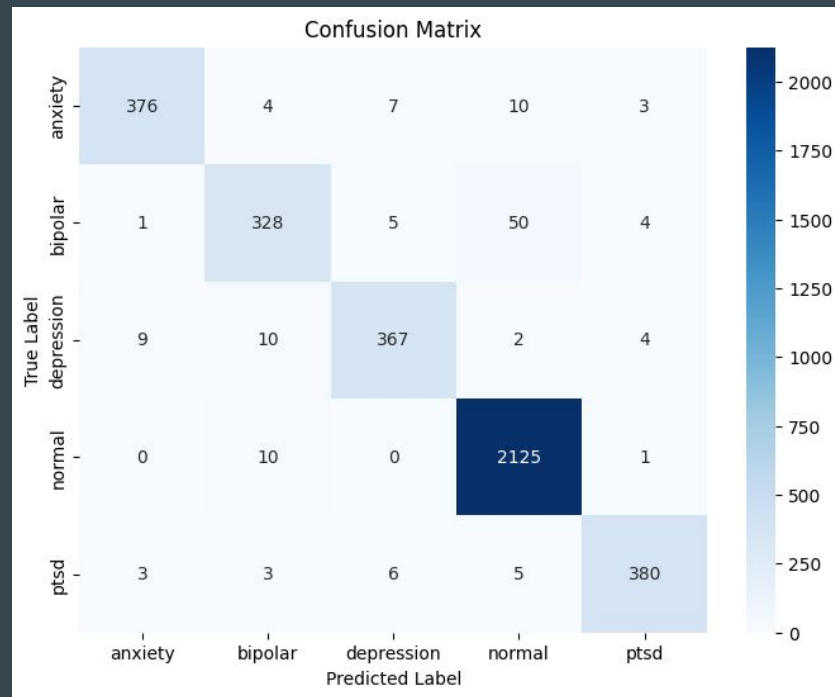


RESULT COMPARISON OF DIFFERENT MODELS AFTER HYPERPARAMETER TUNING

# RESULTS AND ANALYSIS (CONTINUED)



ROC AUC CURVE



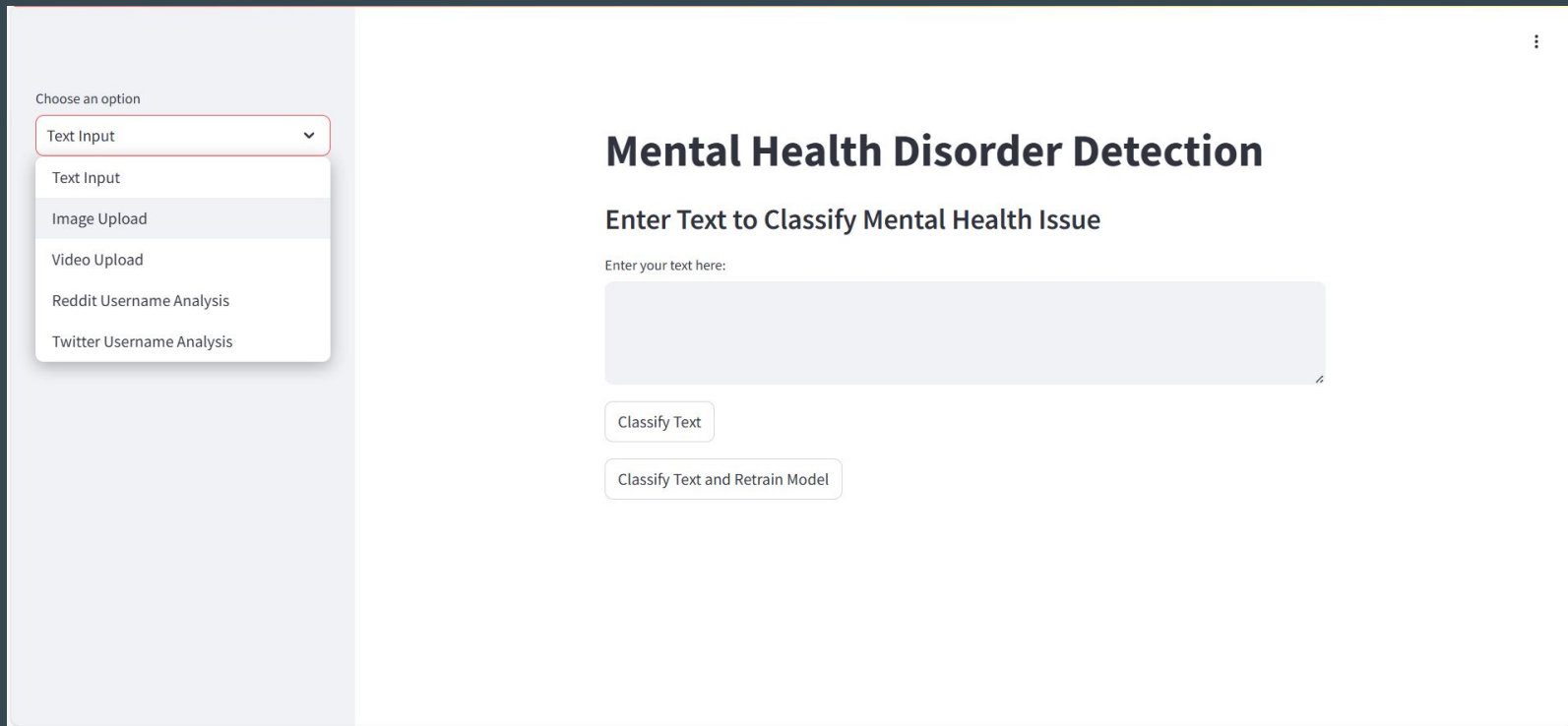
CONFUSION MATRIX

ENSEMBLE MODEL (LOGISTIC REGRESSION + XGBOOST) HAD THE HIGHEST ACCURACY OF 96.09%

# PROTOTYPE

- A functional prototype was developed to classify mental health concerns from social media posts.
- The system allows users to input text, upload images, or submit video for classification.
- Key functionalities include text preprocessing, feature extraction, and prediction using trained machine learning models.
- Real-time results display the most probable mental health concern with confidence scores.
- Allow user to retrain model.
- **Text Classification** : Users can directly input text for immediate analysis and classification.
- **Image-based Classification** : Extracts text from uploaded images using OCR (pytesseract), get captions and facial emotions to classify the content.
- **Video-based Classification** : Processes uploaded video files, extract text from frames, transcribes speech to text, get captions and facial emotions to classify the content.
- **Userprofile Analysis** : Enables analysing user profiles based on posts in Reddit and Twitter

# PROTOTYPE (CONTINUED)



The screenshot displays a web application interface for mental health disorder detection. On the left, a sidebar contains a dropdown menu titled "Choose an option" with a list of options: "Text Input", "Image Upload", "Video Upload", "Reddit Username Analysis", and "Twitter Username Analysis". The "Text Input" option is selected. The main content area features a large heading "Mental Health Disorder Detection" and a subheading "Enter Text to Classify Mental Health Issue". Below the subheading is a text input field with the placeholder text "Enter your text here:". Underneath the input field are two buttons: "Classify Text" and "Classify Text and Retrain Model".

Choose an option

Text Input

Text Input

Image Upload

Video Upload

Reddit Username Analysis

Twitter Username Analysis

## Mental Health Disorder Detection

### Enter Text to Classify Mental Health Issue

Enter your text here:

Classify Text

Classify Text and Retrain Model

## WEB APPLICATION INTERFACE



# PROTOTYPE (CONTINUED)

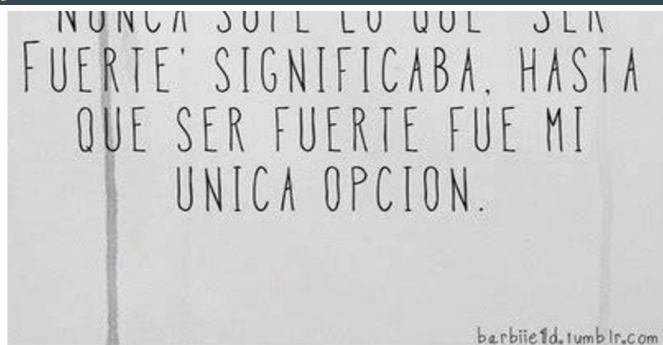
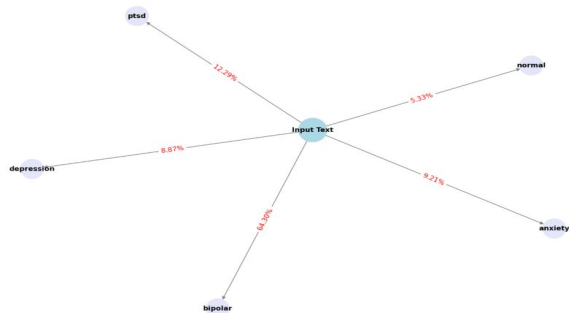
Aggregated Emotion Counts Across All Images



Most Frequent Emotion Across All Images or no Images(Default): sad

Elevated sadness suggests potential depression, while the presence of fear might indicate anxiety or PTSD; seeking professional help for evaluation and exploring coping mechanisms like therapy and mindfulness are crucial.

The most likely mental health concern from all the text obtained is: bipolar with a probability of 64.30%



Uploaded Image

a drawing of a person with a sign on it

Updated Dataset (Last 3 Rows):

	text	mental_health_issue
18,595	Does anybody else get really anxious from hearing a phone ring or even vibrate? I've j	anxiety
18,596	I was born trying to kill myself Came out looking like fucking megamind. Some babie	depression
18,597	When you are alone or feel lonely, many things disappear. Nothing is clear before you	depression

Model is being retrained...

Detailed Status

Retrained Model Accuracy

96.09%

WEB APPLICATION  
INTERFACE

# CONCLUSION

## → *Significance of the Project*

Developed a robust system for early detection of mental health disorders through social media analysis.

## → *Effective Use of Machine Learning*

Leveraged various machine learning models, identifying SVM as the most accurate for sentiment classification.

## → *Impact on Mental Health Awareness*

Provides valuable insights for mental health professionals and public health organizations, enabling proactive interventions.

## → *Potential for Future Development*

Future enhancements with deep learning and multimodal data can lead to even better accuracy and insights.

## → *Commitment to Ethical Practices*

Emphasizes the importance of user privacy and ethical considerations in handling sensitive mental health data.

# WHAT'S NEXT ?

- Explore additional combinations of base models to further optimize the ensemble model's performance.
- Develop a custom transformer model using an embedding encoder tailored for our dataset size and resource constraints.
- Avoid reliance on BERT and its variants due to high resource requirements and lower accuracy on smaller datasets.
- Investigate lightweight transformer architectures for text-based classification to balance accuracy and efficiency.
- Conduct comparative studies on transformer-based versus traditional ML models for text classification accuracy.



# REFERENCES

1. *Choudhury, M. D., De, S., & Counts, S. (2013). Predicting Depression via Social Media. Proceedings of the 7th International Conference on Weblogs and Social Media.*
2. *Guntuku, S. C., Bollen, J., & Lazer, D. (2017). Detecting Mental Illness in Social Media. American Journal of Public Health, 107(8), 1279-1285.*
3. *Mathur, P., Kharat, M., & Patil, S. (2022). Machine Learning Approaches for Mental Health Detection on Social Media: A Systematic Review. IEEE Access, 10, 14376-14388.*
4. *Nadeem, A. (2016). A Study of Depression Identification on Twitter. International Journal of Computer Applications, 141(10), 24-28.*
5. *AlSagri, A., & Ykhlef, M. (2020). A Machine Learning-based Approach for Depression Detection in Social Media. Journal of King Saud University - Computer and Information Sciences, 32(1), 60-66.*
6. *Vaishnavi, A., Rani, A., & Narayan, M. (2022). Application of Machine Learning Algorithms for Mental Health Prediction. International Journal of Data Science and Analytics, 12(2), 175-192.*
7. *Safa, M., Alshahrani, S., & Abunadi, M. (2023). Ethical Considerations in Predicting Mental Health from Social Media: A Roadmap for Future Research. Ethics and Information Technology, 25(1), 15-29.*



THANK YOU