



# GROUP 29

MENTOR : NAIRANJANA CHOWDHURY

## Multimodal AI Framework for Social Media Based Mental Disorder Detection and Personalized Wellbeing Insights

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

PROJECT-III ( PROJ-CS881 ) [ UPID : 008295 ]  
CSE : SEMESTER 8

# CONTENT

1. Motivation
2. Introduction
3. Problem Definition
4. Our Contribution and Uniqueness of Project
5. Research Work
6. Proposed Workflow
7. Implementation
8. Results and Analysis
9. Prototype
10. Additional Features
11. Deployment
12. Conclusion
13. Future Scope
14. 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.

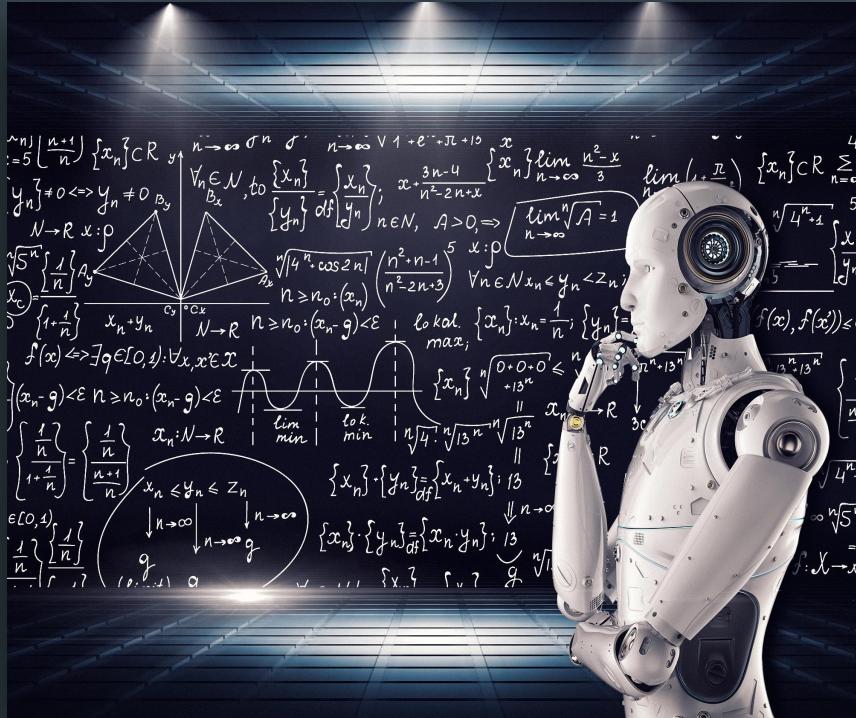
# INTRODUCTION

- **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.



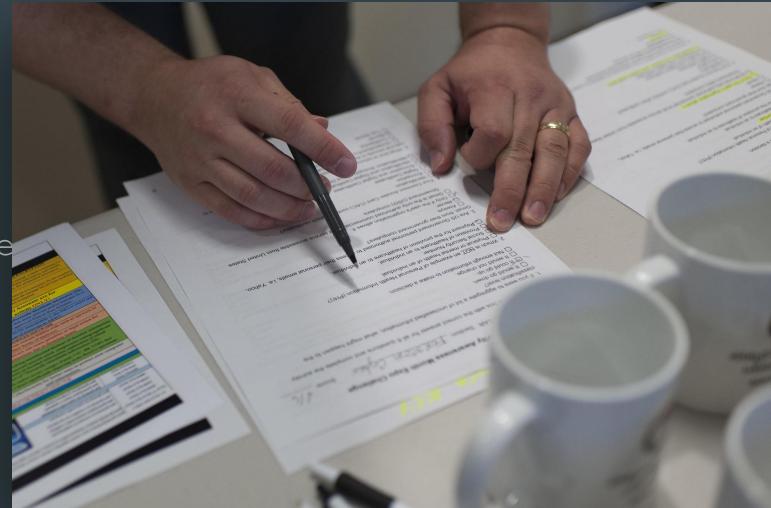
# INTRODUCTION (CONTINUED)

- **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.



# 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.



# OUR CONTRIBUTION AND UNIQUENESS OF PROJECT

- Use technology to assist mental health professionals and enhance public health awareness by implementing different features in a single application
- ❑ **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)*

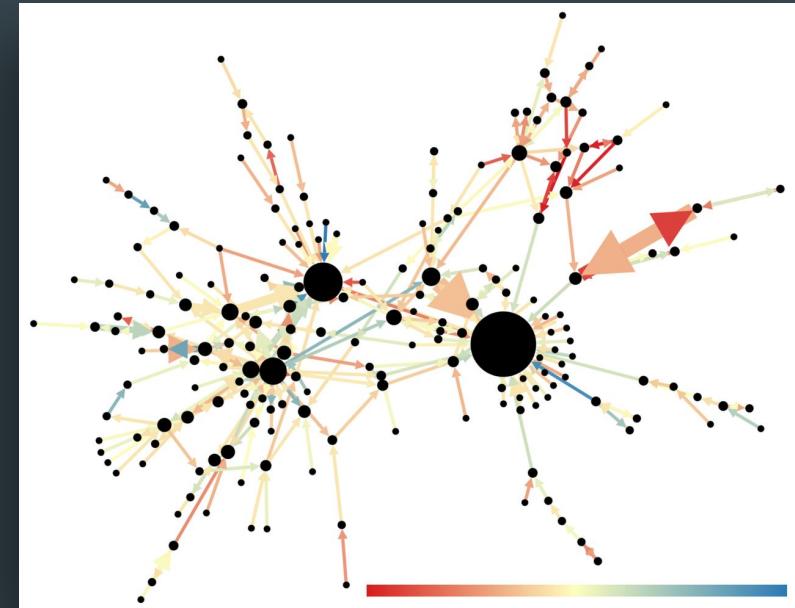
Showed the predictive power of Twitter data in identifying depression through linguistic patterns.

## → *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*

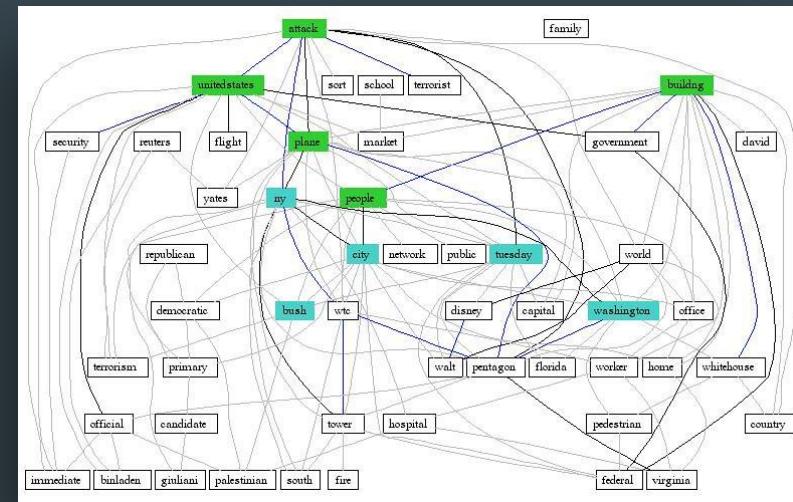
Combined linguistic and behavioral features for more accurate depression detection on Twitter.

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

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

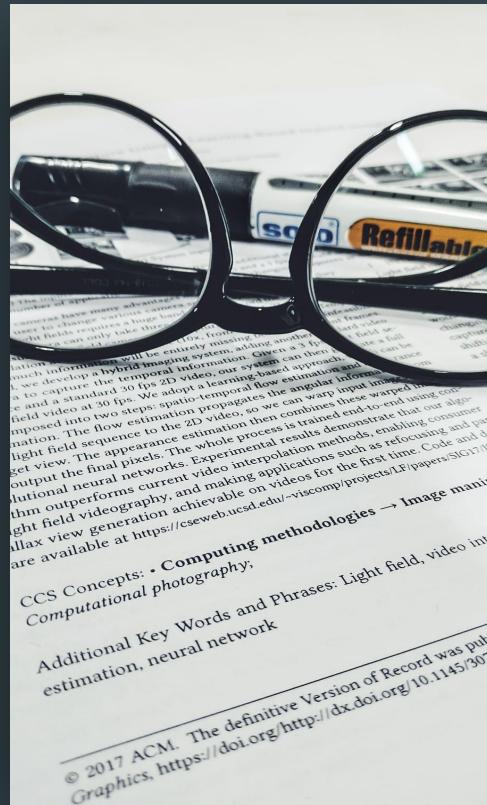
## → *Ethical considerations by Safa et al. (2023)*

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



# RESEARCH WORK (CONTINUED)

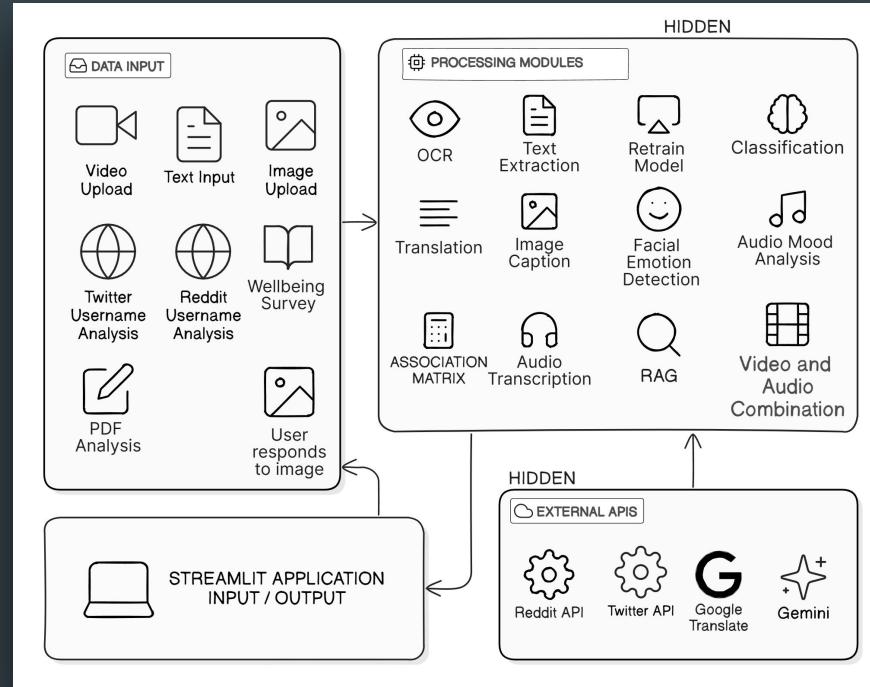
- The study "**Single classifier vs. ensemble machine learning approaches for mental health prediction**" demonstrates that Gradient Boosting (88.80%) and Neural Networks (88.00%) outperform ensemble models (85.60%) for mental health issue prediction, based on survey data from OSMI.
- The paper "**Ensemble of hybrid model-based technique for early detection of depression**" introduces a hybrid SVM-MLP model with SMOTE for class balancing, achieving 99.39% accuracy and 99.51% F1-score, highlighting its efficacy in early depression detection.
- The research "**Survey of transformers and towards ensemble learning using transformers for NLP**" reviews transformer models and explores ensemble learning with transformers, showcasing their superior performance in NLP tasks like sentiment analysis and text generation.



# PROPOSED WORKFLOW

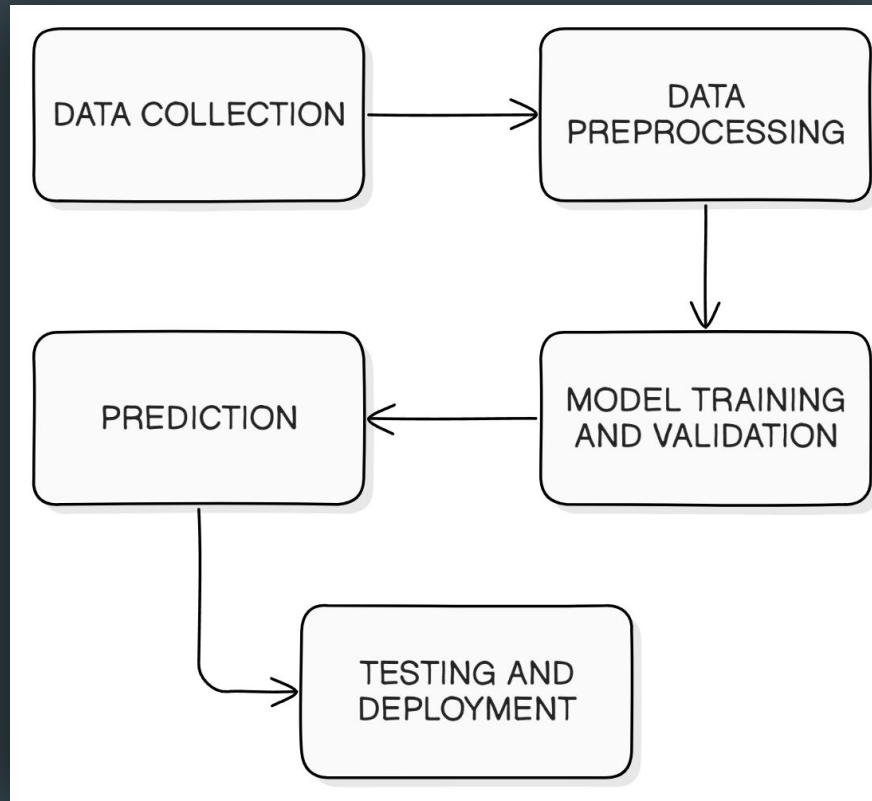
- **Data Collection** : Collect Reddit posts using the PRAW API and extract labeled data for mental health categories like Normal, Anxiety, Depression, PTSD, and Bipolar.
- **Data Preprocessing** : Clean the collected data by removing URLs, stopwords, and special characters. Normalize and tokenize the text, followed by converting it into numerical formats ( *TF-IDF, Bag Of Words, Word2Vec, LIWC, N-Gram* ) for analysis.
- **Model Training and Evaluation** : Train machine learning models including Logistic Regression, Naive Bayes, SVM, Random Forest, XGBoost, KNN, LSTM and Transformers. 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 multiple 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)



PROJECT MODULES

# IMPLEMENTATION

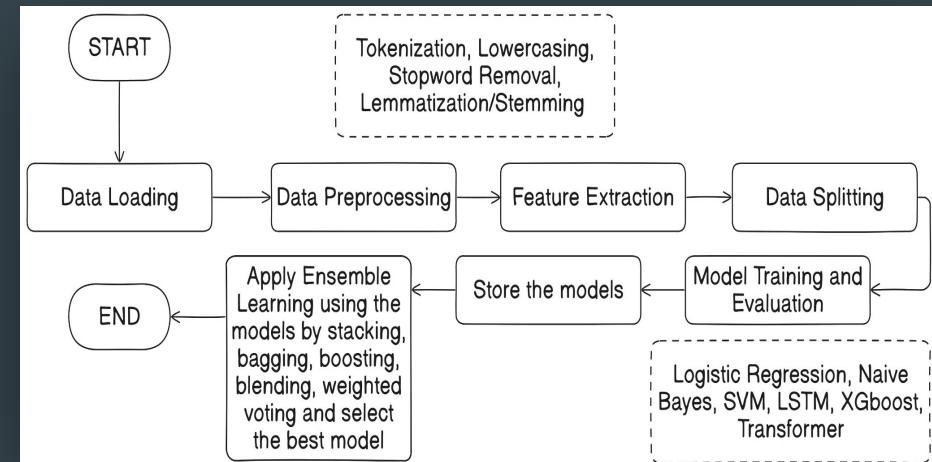
- **Data Collection** : Scrapped Reddit posts related to mental health issues and for analysis.
- **Data Preprocessing** : Cleaned and normalized the text by removing URLs, stop-words, punctuation, and applied tokenization and lemmatization techniques.
- **Feature Extraction** : Utilized Bag of Words, Term Frequency-Inverse Document Frequency (TF-IDF), Word2Vec, LIWC, N-Gram 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. Also applied Stratified K-fold Cross Validation

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
FR-005	Text Analysis	Completed
FR-006	Image Upload Analysis	Completed
FR-007	Video Upload Analysis	Completed
FR-008	PDF Upload Analysis	Completed
FR-009	User response to image	Completed
FR-010	Reddit and Twitter Username Analysis	Completed
FR-011	Wellbeing survey and mapping using association matrix	Completed
FR-012	Application Deployment and Model Retraining	Completed
NFR-001	Scalability and Performance	Completed

## FUNCTIONAL REQUIREMENTS

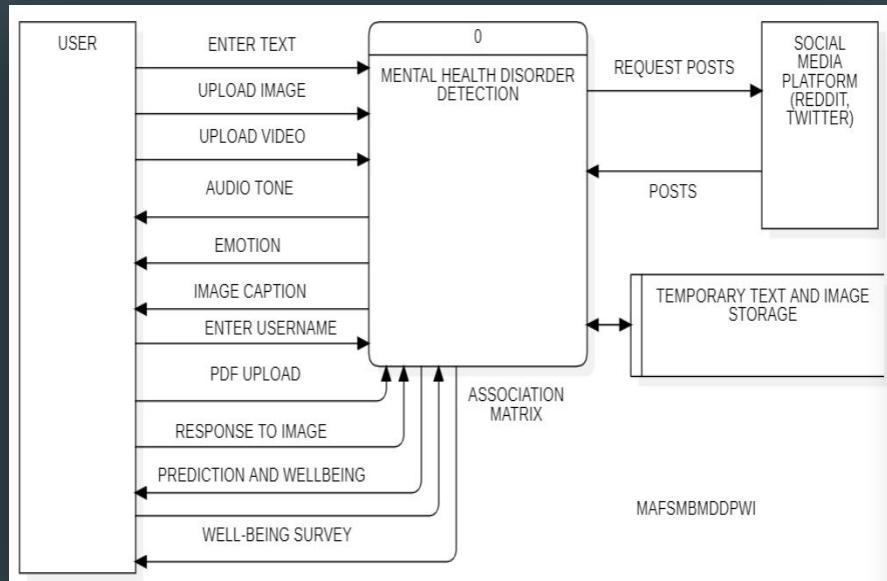
# IMPLEMENTATION (CONTINUED)

- **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, Transformer and Ensemble Models.
- **Hyperparameter Tuning** : Applied RandomizedSearchCV to optimize the performance of some models.
- **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 and deployed using Google Colab, Pyngrok and Streamlit.



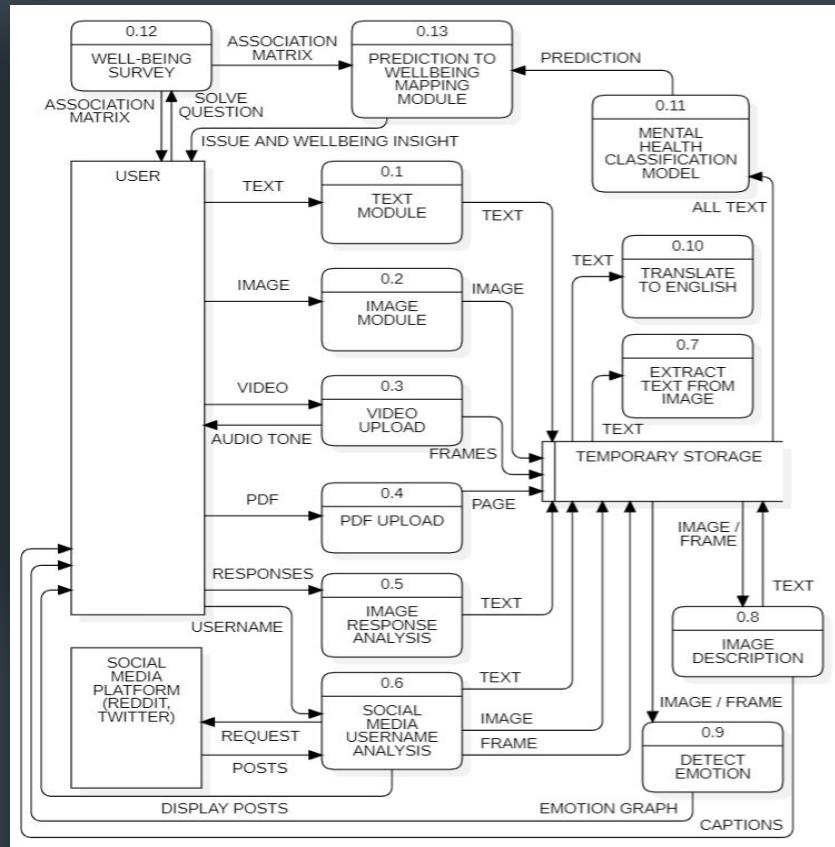
MODEL WORKFLOW

# IMPLEMENTATION (CONTINUED)

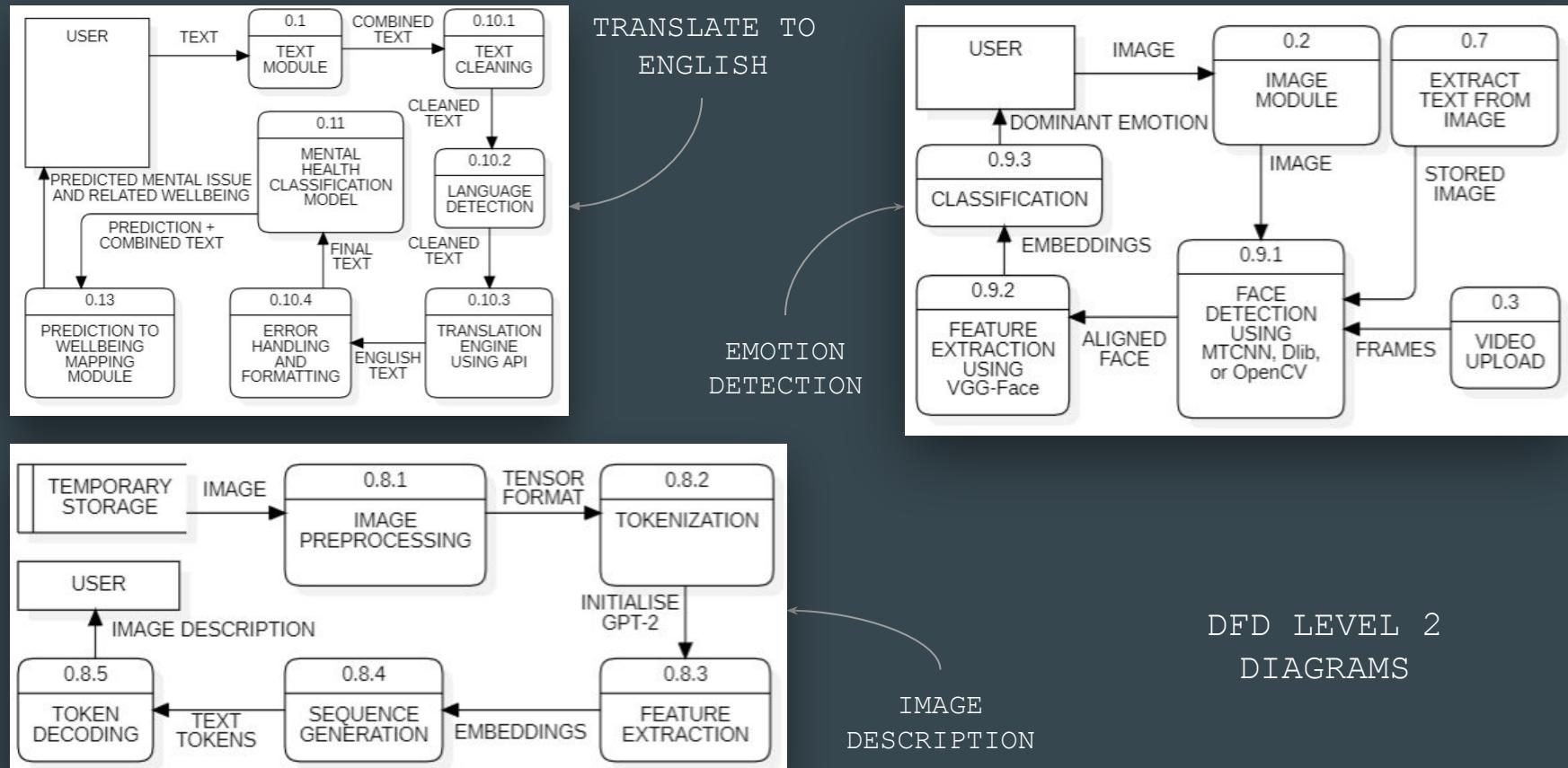


DFD LEVEL 0

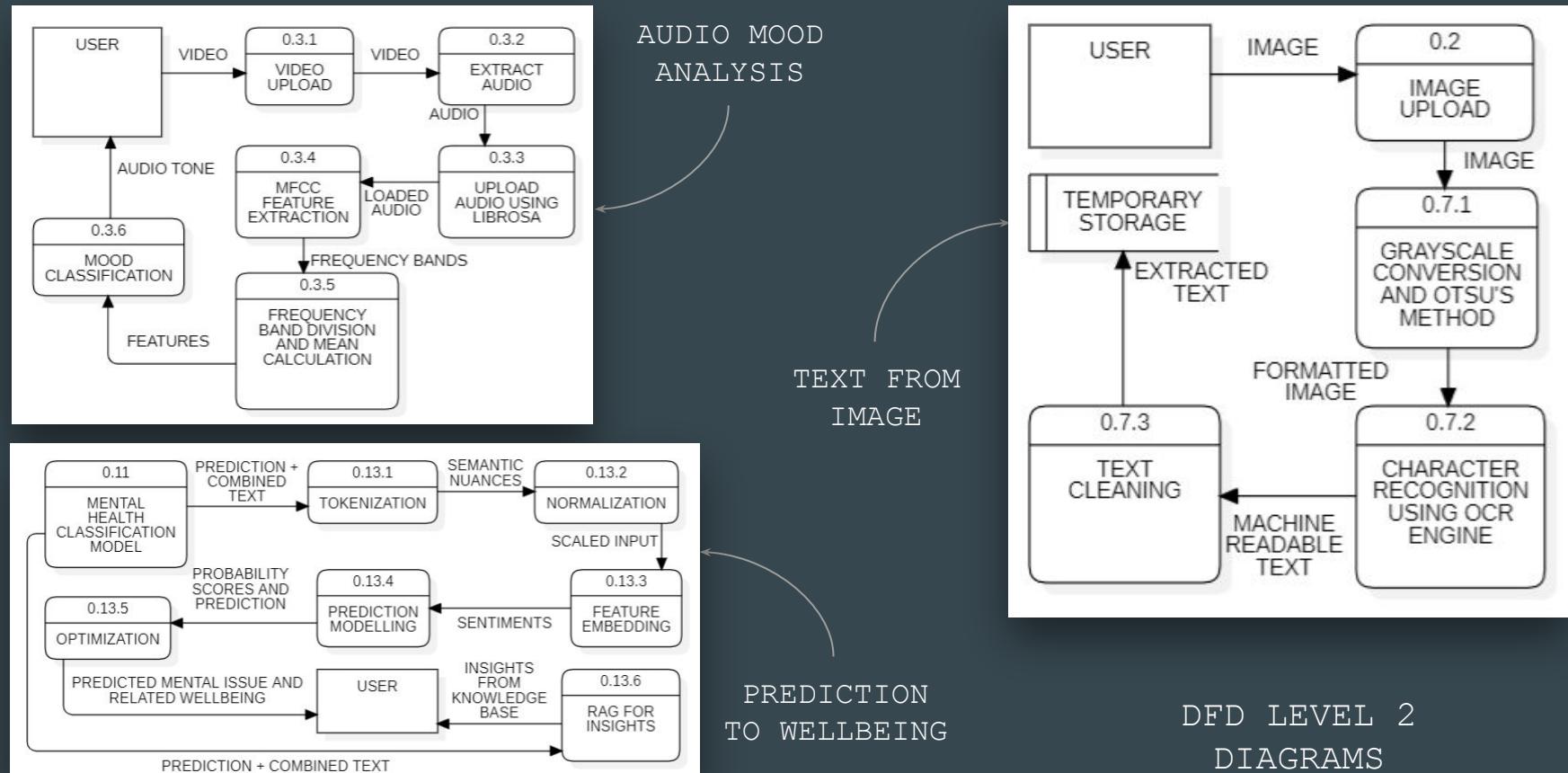
DFD LEVEL 1



# IMPLEMENTATION (CONTINUED)



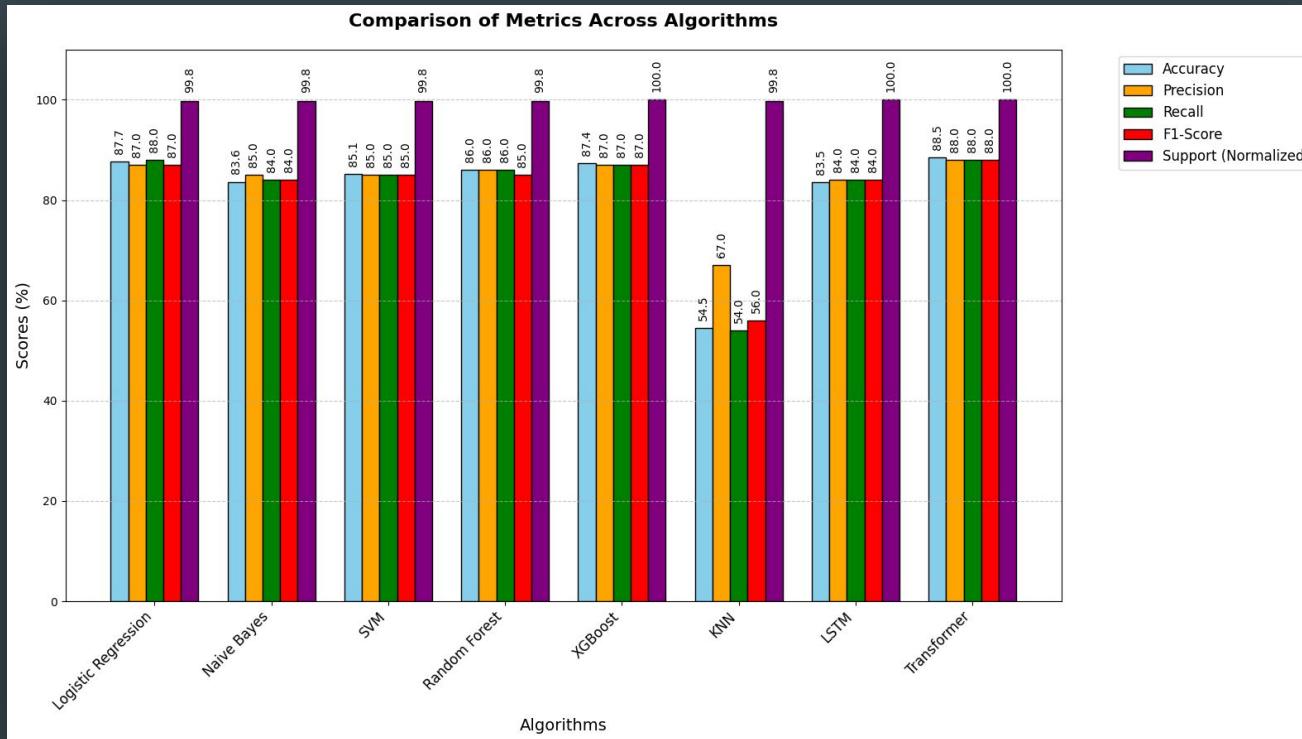
# IMPLEMENTATION (CONTINUED)



# RESULTS AND ANALYSIS

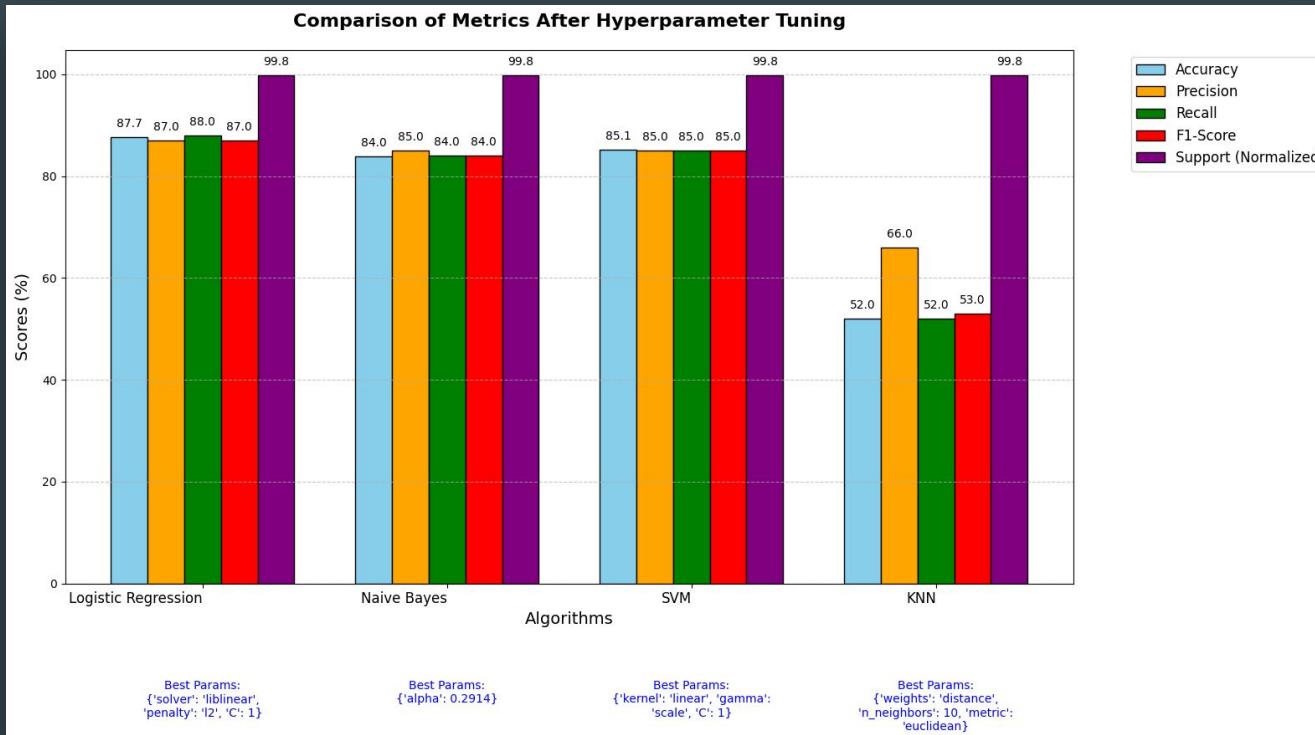
- **Logistic Regression** : Achieved consistent performance with high accuracy. Precision and recall indicate reliable classification for balanced datasets.
- **Naive Bayes** : Performed well with text data, especially for independent features. Precision was slightly lower for imbalanced classes but remained effective overall. Got better performance after applying hyperparameter tuning.
- **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.
- **Transformer** : Gave the highest accuracy when implemented separately. Used in Ensemble Learning and improved the overall accuracy of the final model.
- **K-nearest Neighbours** : Performed poorly and gave the worst accuracies among all.

# RESULTS AND ANALYSIS (CONTINUED)



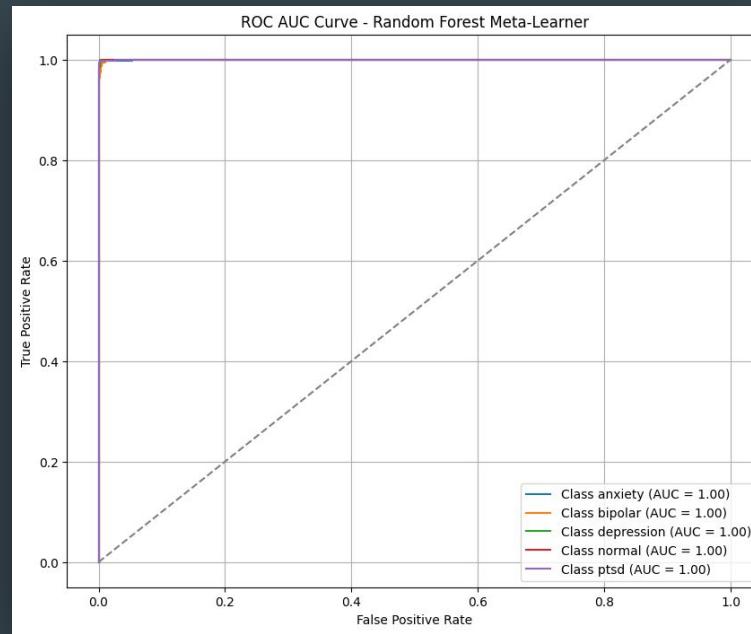
RESULT COMPARISON OF DIFFERENT MODELS

# RESULTS AND ANALYSIS (CONTINUED)

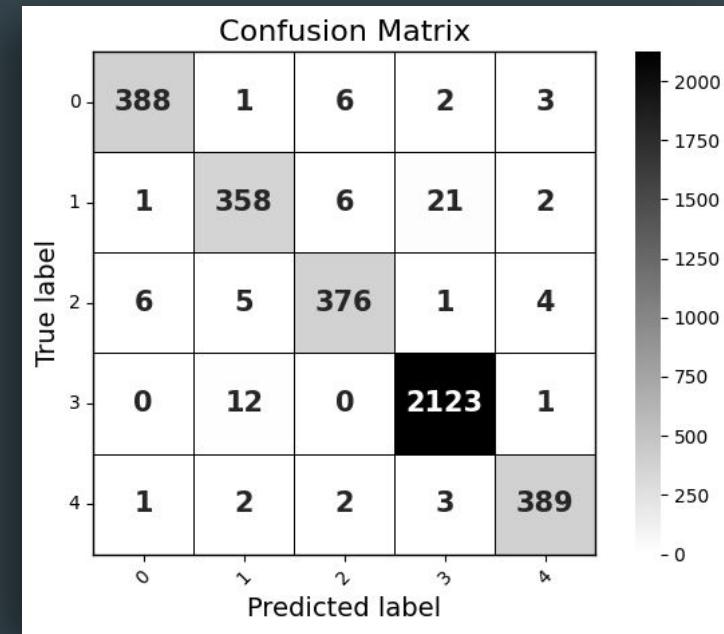


RESULT COMPARISON OF DIFFERENT MODELS AFTER HYPERPARAMETER TUNING

# RESULTS AND ANALYSIS (CONTINUED)



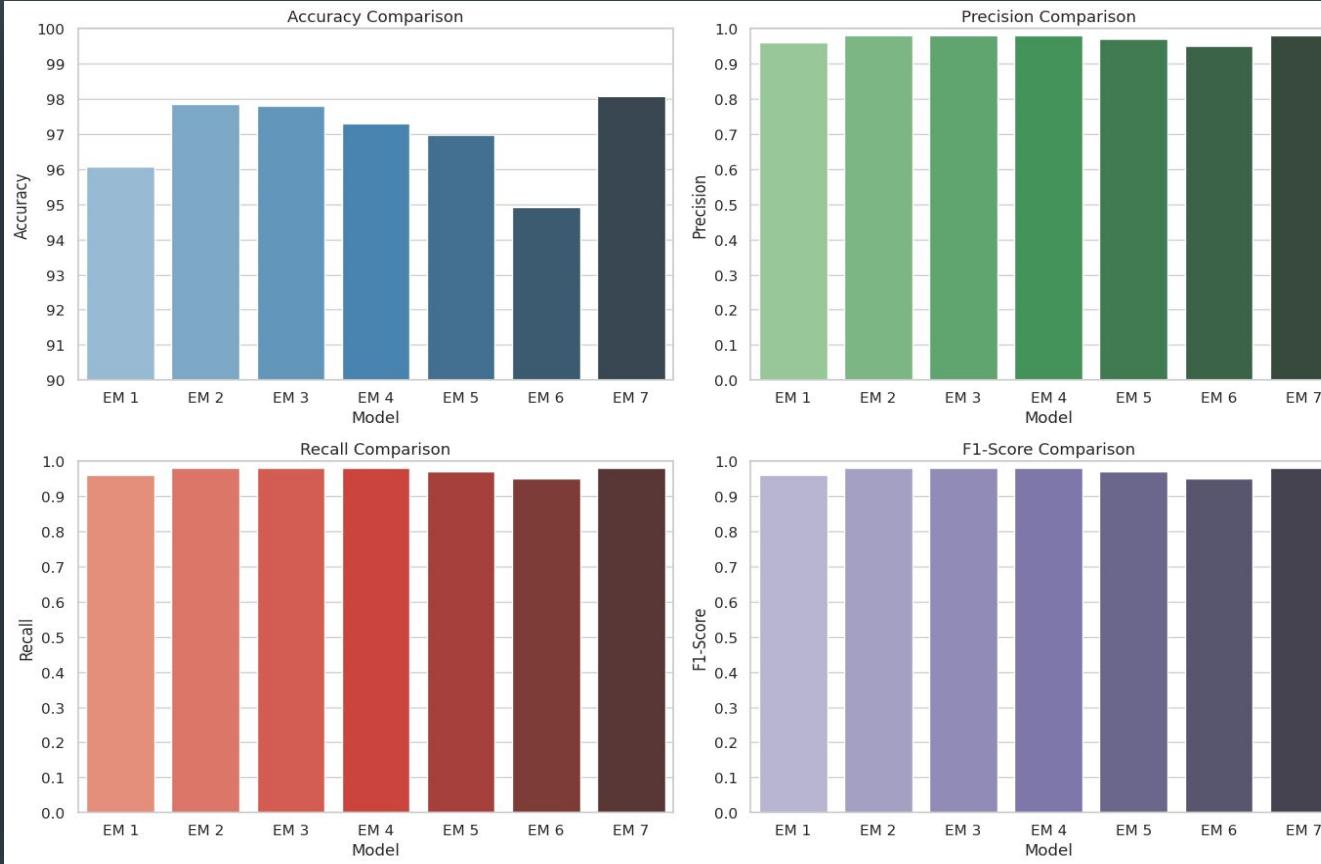
ROC AUC CURVE



CONFUSION MATRIX

ENSEMBLE MODEL (BASE MODELS : LOGISTIC REGRESSION, NAIVE BAYES, SVM, XGBOOST, LSTM, TRANSFORMER & META LEARNER : RANDOM FOREST) WITH ACCURACY OF 98.03% WAS USED IN WEB APPLICATION

# RESULTS AND ANALYSIS (CONTINUED)



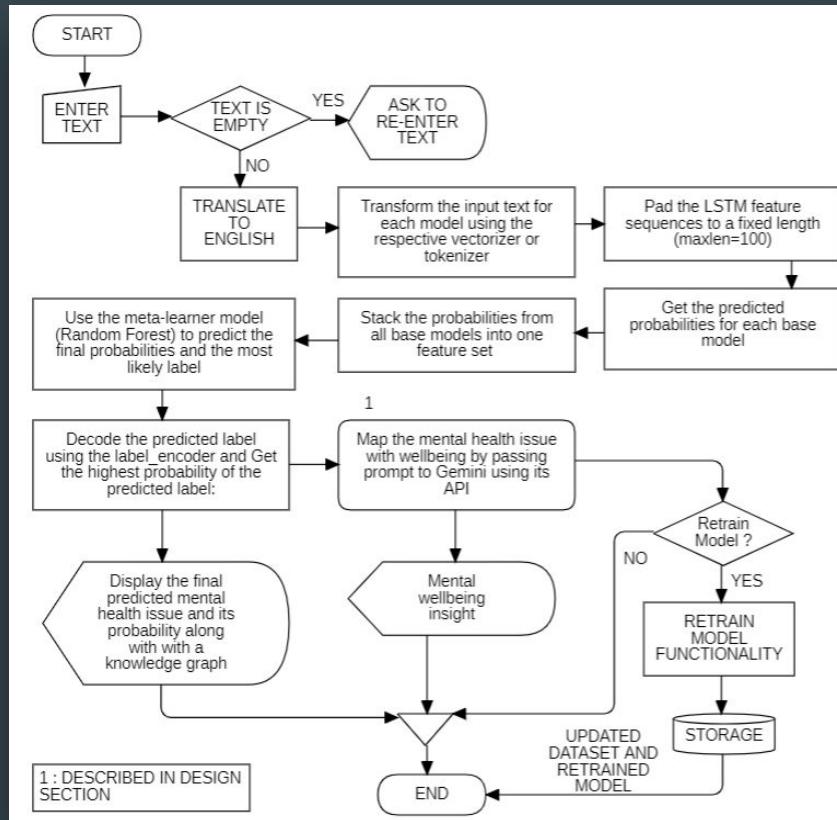
COMPARISON OF  
DIFFERENT ENSEMBLE  
MODELS

# 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)



TEXT CLASSIFICATION

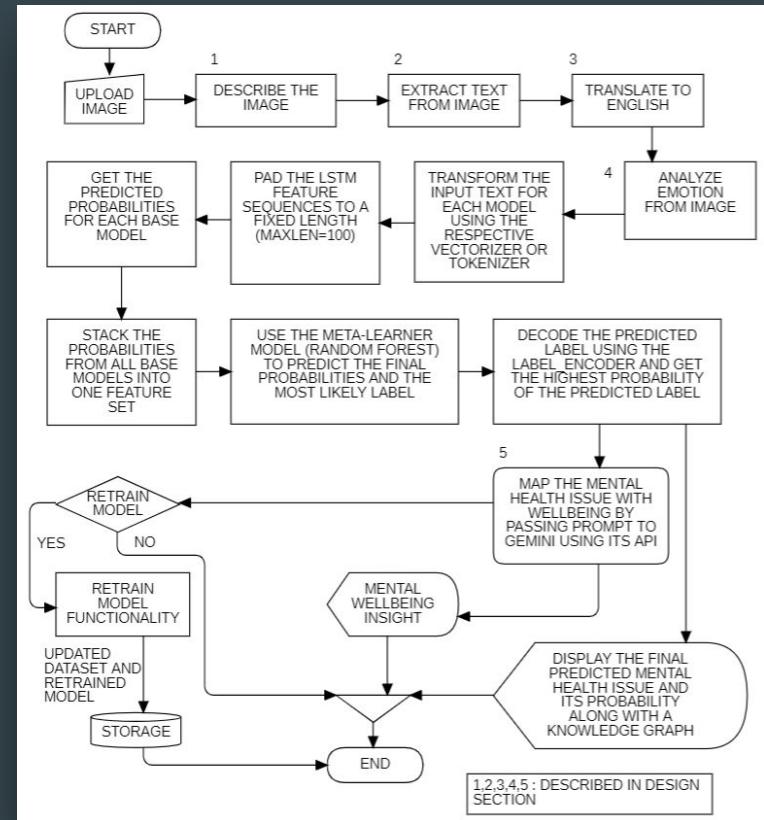
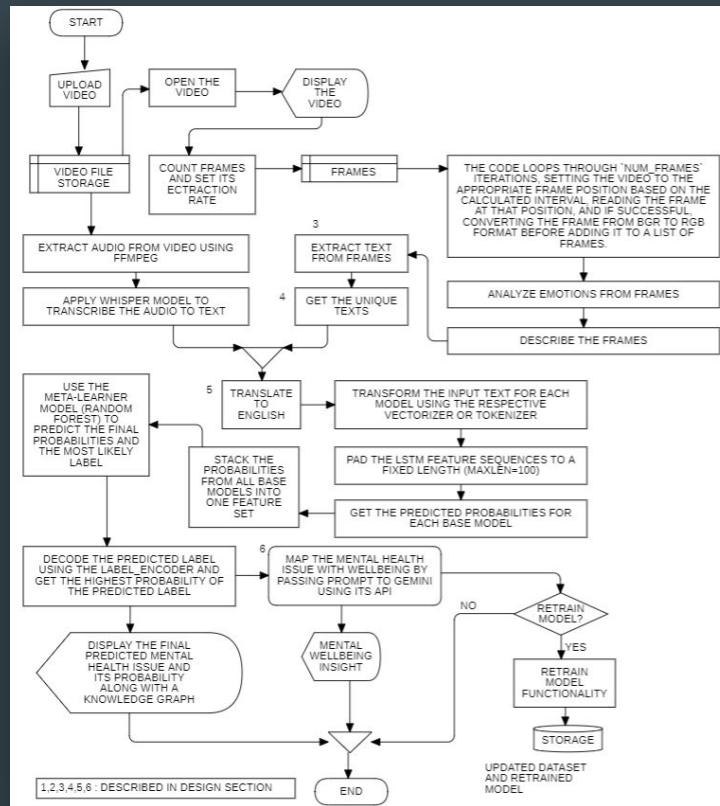
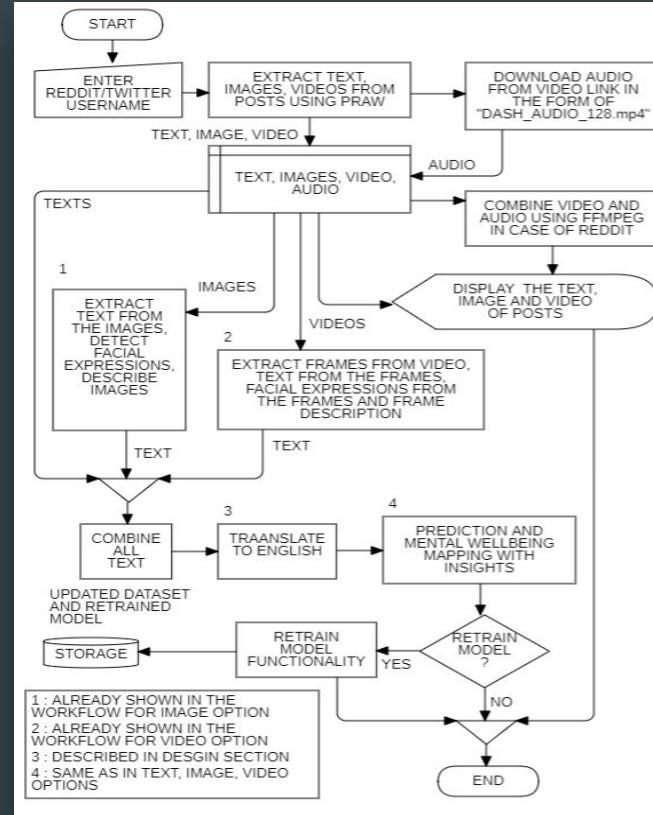


IMAGE CLASSIFICATION

# PROTOTYPE (CONTINUED)



VIDEO ANALYSIS



REDDIT AND TWITTER USERNAME ANALYSIS

# PROTOTYPE (CONTINUED)

The screenshot shows a web application interface titled "Mental Health Disorder Detection". On the left, there is a sidebar with a dropdown menu titled "Choose an option" containing the following options: Text Input (selected), Image Upload, Video Upload, PDF Upload, Responses to Image, Reddit Username Analysis, Twitter Username Analysis (disabled), and Well-being Survey. The main content area has a title "Mental Health Disorder Detection" and a subtitle "Enter Text to Classify Mental Health Issue". Below the subtitle is a text input field with placeholder text "Enter your text here:". To the right of the input field are two buttons: "Classify Text" and "Classify Text and Retrain Model".

WEB APPLICATION INTERFACE

# PROTOTYPE (CONTINUED)

## Mental Health Disorder Detection

### Enter Text to Classify Mental Health Issue

Enter your text here:

অসুস্থির কাছে হেলেগে হেলে। সামাজিকের এক ধরনের আইনিরতা অনুভূত কর, যুক্তি মধ্যে চাপ বাড়তে থাকে, খুব নিতে কষ করা তুমি আমা, এটি কেবল মনসময়ের ভুল, কিন্তু তাপমাত্র এটি দেখে দেখে প্রতিটি কথে উচ্চিতার পড়ে। শুধুমাত্র এক অক্ষুণ্ণ গহনের মধ্যে অনুভূত হয়, যেখানে তুমি একা এবং অসহ্য।

Classify Text

Classify Text and Retrain Model

### Translated Text (to English):

Your mind seems to be stuck in a thousand thoughts and doubts every day. A feeling like a nerve-wracking, a kind of pressure that never goes away. An unknown fear works inside you, which says —“Something bad will happen.” Maybe you know, nothing is likely to happen, but deep down in your mind an impossible fear awakens. A kind of restlessness is felt throughout your body, the pressure in your chest increases, it becomes difficult to breathe. You know, this is just a mistake in concentration, but still it seems to spread to every cell of your body. The world feels like a dark hole, where you are alone and

chest increases, it becomes difficult to breathe. You know, this is just a mistake in concentration, but still it seems to spread to every cell of your body. The world feels like a dark hole, where you are alone and helpless.

The most likely mental health concern from all the text obtained is: anxiety with a probability of 99.95%

### Wellbeing Insight:

**1. Autonomy and Anxiety:** Anxiety can severely impair autonomy. The overwhelming fear and worry associated with anxiety can make it difficult for individuals to confidently assert their opinions or make independent decisions, even when they disagree with the majority ("I have confidence in my opinions, even if they are contrary to the general consensus"). The need for external reassurance and validation becomes prominent, hindering self-regulation and leading to dependence on others for decision-making. This dependence contradicts the Ryff scale's definition of autonomy.

**2. Environmental Mastery and Anxiety:** Anxiety significantly diminishes environmental mastery. The inability to manage overwhelming feelings can lead to avoidance behaviors, preventing individuals from effectively managing daily tasks and opportunities ("In general, I feel I am in charge of the situation in which I live"). Anxiety can manifest as procrastination, difficulty focusing, and an inability to anticipate and plan for future events, resulting in a decreased sense of control over one's environment.

**3. Personal Growth and Anxiety:** Anxiety often impedes personal growth. The constant worry and fear can stifle exploration of new experiences and opportunities ("I think it is important to have new experiences that challenge how you think about yourself and the world"). Individuals might avoid situations that

## Mental Health Disorder Detection

### Upload an Image to Extract and Classify Text

Upload an Image

Drag and drop file here  
Limit 200MB per file • JPEG, PNG, WEBP, BMP, TIFF, TIF

OIP (2).jpg 26.7KB

[Browse files](#)



## Mental Health Disorder Detection

### Upload a Video to Extract and Classify Text

Choose a video file

Drag and drop file here  
Limit 200MB per file • MP4, MOV, AVI, MPEG

smp3.mp4 3.7MB



## Mental Health Disorder Detection

### Enter Reddit Username for Analysis

Enter Reddit username:

flowerpower0601

Analyze

Recent Text Posts:

```
[{"id": "t1_0", "text": "Taylor Swift's 'Eras' show. What's ACTUALLY going on? What do you guys think of this?", "author": "flowerpower0601"}, {"id": "t1_1", "text": "Taylor Swift's 'Eras' show. What's ACTUALLY going on? [removed]", "author": "flowerpower0601"}, {"id": "t1_2", "text": "Taylor Swift und die Eras Tour", "author": "flowerpower0601"}]
```

Ich habe letztens dieses Video gefunden, welches behauptet das die Eras Tour von Taylor Swift überhaupt nicht live ist und das selbst, dass was sich "live" anhört einfach pre-recorded ist und die Band gar nicht live spielt sondern die ganze Show ein Backing Track ist, der abgespielt wird.

## Mental Health Disorder Detection

### Enter Twitter Username for Analysis

Enter Twitter username:

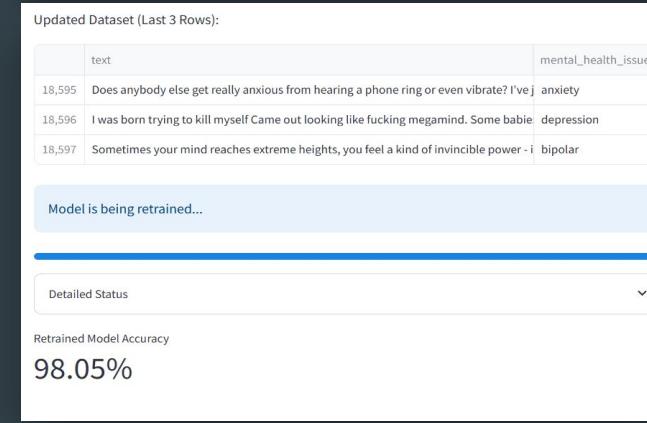
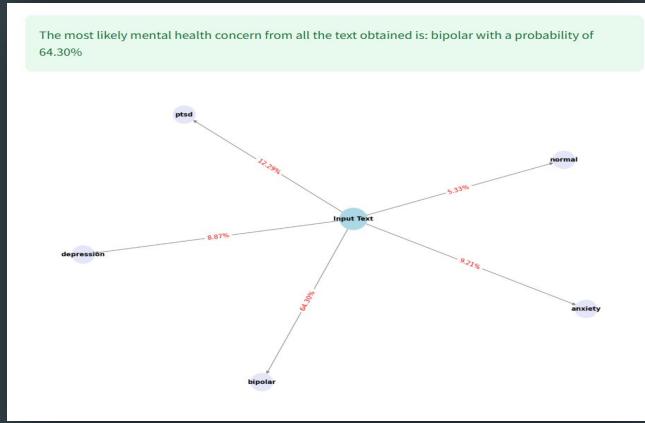
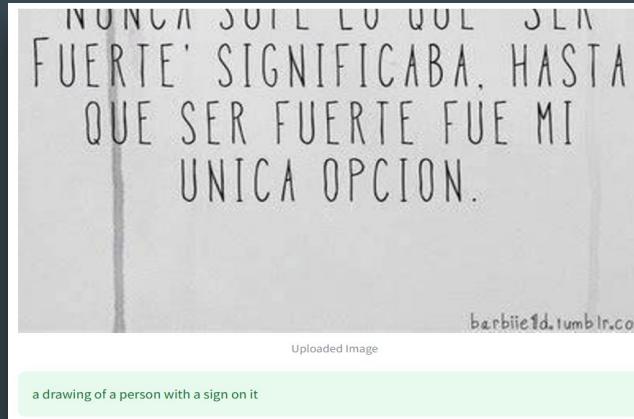
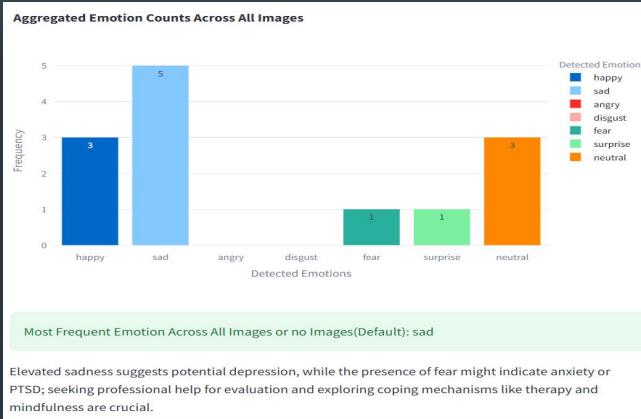
narendramodi

Analyze

Recent Text Posts from Tweets:

```
[{"id": "t1_0", "text": "Well said. It is good that this truth is coming out, and that too in a way common people can see it.", "author": "narendramodi"}, {"id": "t1_1", "text": "A fake narrative can persist only for a limited period of time. Eventually, the facts will always come out! https://t.co/8Xx0s5Qqe2y", "author": "narendramodi"}, {"id": "t1_2", "text": "महान बालाशुभ्र ठाकरे जी यथा पूर्णतयामिनि मी लानं आदर्शजली अर्पण करते, महाराष्ट्रव विकास अग्नि मराठी लोकाचा सुखाता यासाठी अप्राप्ती असे ते एक द्रष्टे याक्षित लाते. भारतीय संस्कृती अग्नि मूर्त्युचे सत्परेन कर्त्तव्य यावतीचा अभियान युद्धात करण्यातर लांबा द... https://t.co/oHRv3DUFKj", "author": "narendramodi"}]
```

# PROTOTYPE (CONTINUED)



## WEB APPLICATION INTERFACE

# ADDITIONAL FEATURES

## PDF UPLOAD ANALYSIS

### Automated Text Extraction:

- Utilize OCR to extract text from uploaded PDF files.
- Convert scanned documents into machine-readable text.

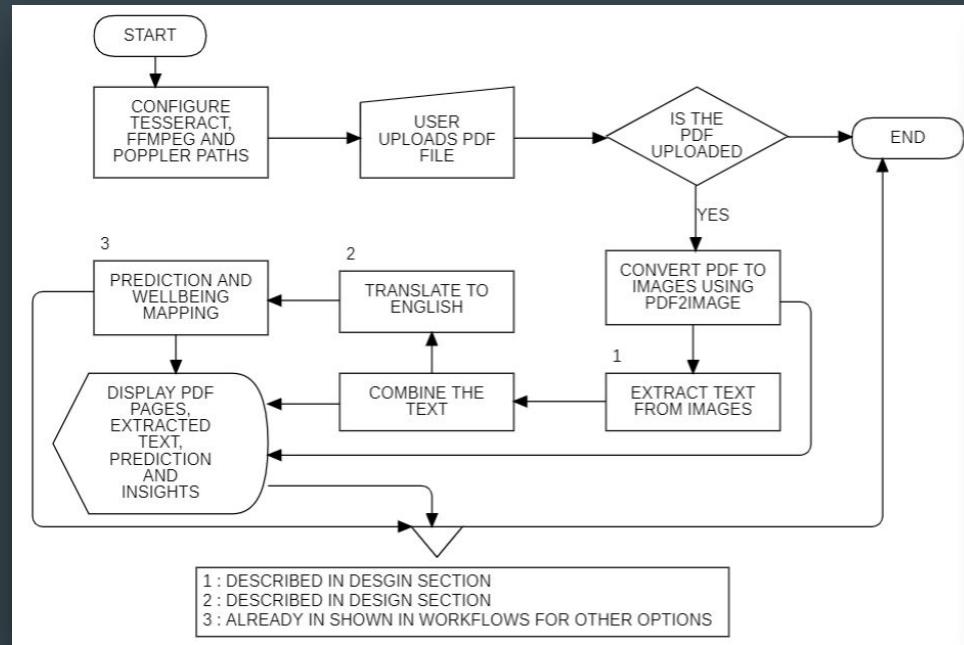
### Data Enrichment:

- Integrate PDF content into the mental health dataset.
- Enhance analysis with diverse sources of user-generated content.

### Subsequent Analysis:

- Apply models to evaluate and classify the extracted text.
- Generate insights from PDF-based textual data.

# ADDITIONAL FEATURES (CONTINUED)



PDF UPLOAD OPTION

## Mental Health Disorder Detection

Upload a PDF to Extract and Classify Text

Upload a PDF

Drag and drop file here  
Limit 200MB per file • PDF

Handwritten text.pdf 1.5MB X

Extracted Pages

I don't know where to begin. Maybe it doesn't even matter. Lately, everything feels like a blur, like I'm moving through life in slow motion while the rest of the world rushes past me. It's exhausting. Waking up every day, knowing it's going to be the same as yesterday—the same weight pressing down on my chest, the same thoughts circling in my head, the same emptiness that never seems to go away.

# ADDITIONAL FEATURES (CONTINUED)

## USER RESPONSE TO IMAGE SHOWN

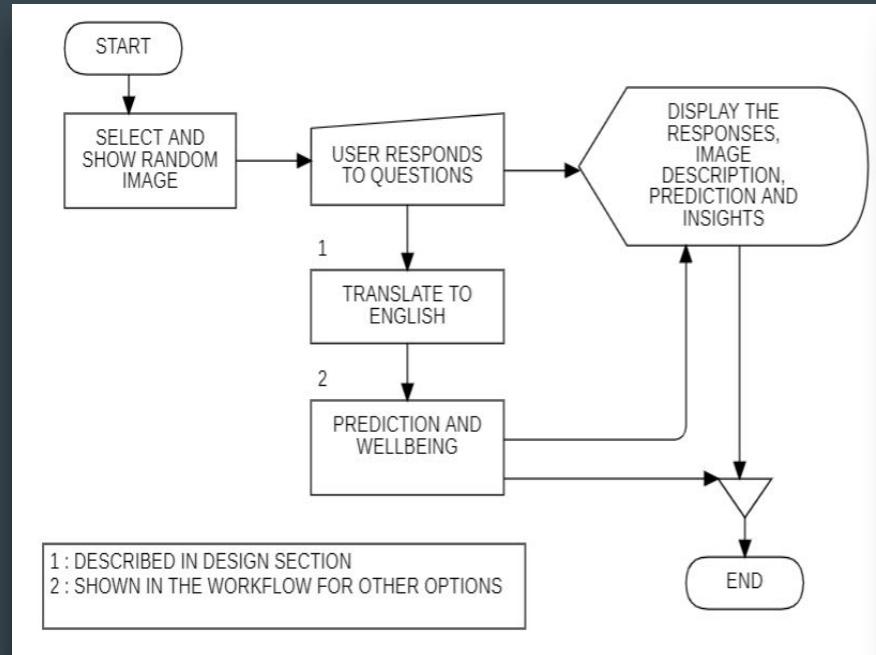
- **Interactive Image-Based Feedback:**
  - ◆ Capture user responses to images (e.g., via rating open-ended input) for some questions similar to Rorschach Inkblot Test.
- **Enhanced User Engagement:**
  - ◆ Integrate visual feedback to improve overall classification accuracy.
  - ◆ Personalize insights based on user responses to visual stimuli.

### Mental Health Disorder Detection

Describe Image and Classify Responses



# ADDITIONAL FEATURES (CONTINUED)



## Answer the following questions:

What do you see in this image?

I see a woman with a distant, haunted expression. Her head seems to be bursting with chaotic, mechanical fragments, almost like intrusive thoughts or painful memories that she cannot escape. The background is dark and fragmented, with shadowy figures that seem distant yet significant, as if representing people from her past or unresolved trauma.

What emotions does this image evoke in you?

It evokes feelings of anxiety, isolation, and distress. The mechanical structures in her mind feel heavy, symbolizing an overwhelming mental burden. The color palette, with its contrast between dark blues and fiery oranges, makes me feel trapped between the coldness of detachment and the burning intensity of past trauma.

Does this image remind you of anything from your past?

Yes, it reminds me of moments where my thoughts felt uncontrollable, where memories of past pain resurfaced without warning. The figures in the distance mirror the feeling of being surrounded by others yet feeling completely alone. The construction-like elements in the background resemble a city lost in time, much like certain places tied to difficult memories.

## USER RESPONSE TO IMAGE OPTION

# ADDITIONAL FEATURES (CONTINUED)

## WELLBEING SURVEY

### Comprehensive Survey Deployment:

- Collect user responses on mental health and well-being using a detailed survey.
- Incorporate questions based on the Ryff Scale (Self-Acceptance, Autonomy, etc.).

### Dynamic Association Matrix:

- Update the association matrix using survey response data.
- Compute weighted scores and similarity metrics (e.g., weighted sum, cosine similarity, euclidean distance).

### Personalized Well-being Insights:

- Generate targeted recommendations based on the computed association matrix.
- Display consensus results that inform users about key well-being dimensions.

# ADDITIONAL FEATURES (CONTINUED)

Questions with (R) are reversed scored.

1 → Strongly Disagree  
2 → Disagree  
3 → Slightly Disagree  
4 → Slightly Agree  
5 → Agree  
6 → Strongly Agree

**Q00. What is Your Predicted Mental Issue?**

Anxiety  
 Bipolar  
 Depression  
 Normal  
 PTSD

**Q01. When I look at the story of my life, I am pleased with how things have turned out.**

1  2  3  4  5  6

**Q02. In many ways I feel disappointed about my achievements in life. (R)**

Responses Submitted Successfully!

**Overall Scores (Max: 12 for each parameter) and Interpretation :**

**Self Acceptance: 10 (High Scorer)**

Possesses a positive attitude toward the self; acknowledges and accepts multiple aspects of self, including good and bad qualities; feels positive about past life.

**Positive Relations with Others: 10 (High Scorer)**

Has warm, satisfying, trusting relationships; concerned about the welfare of others; capable of strong empathy, affection, and intimacy.

**Autonomy: 10 (High Scorer)**

Is self-determining and independent; able to resist social pressures; regulates behavior from within and follows personal standards.

**Environmental Mastery: 10 (High Scorer)**

Has a sense of control over one's environment; able to handle challenges effectively; feels competent in various situations.

**Specific Parameter Based Insight:**

A person struggling with depression often finds it challenging to maintain positive relationships and cultivate autonomy. However, focusing on these areas within the Ryff scale can significantly contribute to recovery. Here's practical advice, keeping in mind that progress is gradual and setbacks are normal:

**Improving Positive Relations with Others:**

- Start small, don't overwhelm: Instead of aiming for large social gatherings, start with small, manageable interactions. A brief phone call with a trusted friend or a short coffee date is a good beginning. Focus on the quality of connection, not quantity.
- Identify supportive individuals: Recognize and prioritize connections with people who are understanding, patient, and supportive. They can provide a safe space to express feelings without judgment. Limit contact with those who might exacerbate your depression.
- Communicate honesty (but cautiously): Sharing your struggles with trusted individuals can be incredibly helpful, but do so at your own pace. Start with small disclosures and gauge their reactions. If you feel judged or invalidated, protect your emotional well-being and limit further disclosure.
- Practice active listening: Focusing on others helps shift attention away from internal struggles. Truly listening to friends and family strengthens bonds and fosters reciprocal support.
- Engage in shared activities: Participate in activities you enjoy with others, even if it's just a walk in the park or watching a movie. Shared experiences create connection and positive memories.
- Join a support group: Connecting with others who understand depression can reduce feelings of isolation and provide a sense of community. Online support groups can offer accessibility if in-person meetings are challenging.
- Accept help: Don't hesitate to ask for help when needed. This could range from practical assistance

parameter	anxiety	bipolar	depression	normal	ptsd
self acceptance	3	2	1	5	2
positive relations with others	3	4	2	5	3
autonomy	2	3	2	5	1
environmental mastery	2	4	1	5	4
purpose in life	4	4	1	5	3
personal growth	3	3	1	5	4

Total Number of Respondents (2025-02-08): 1

Updated Association Matrix successfully.

**Updated Association Matrix:**

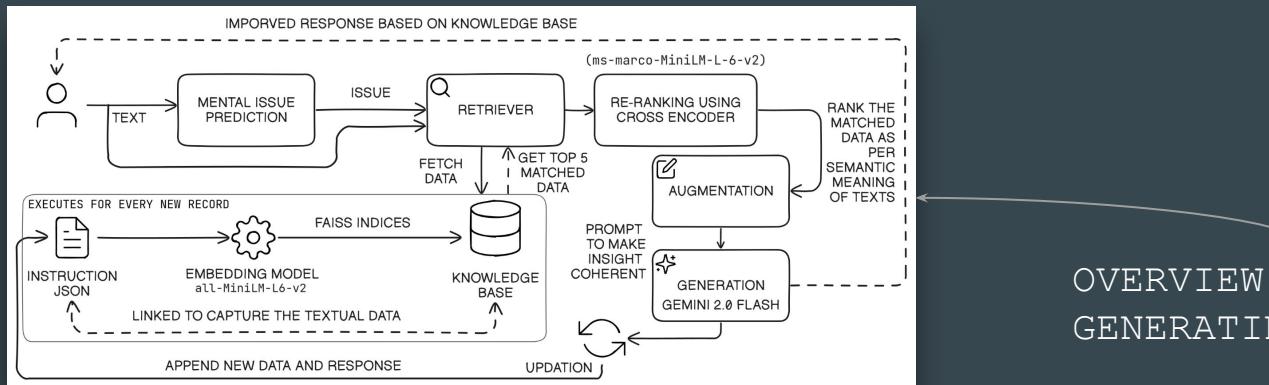
parameter	anxiety	bipolar	depression	normal	ptsd
self acceptance	3	2	1	5	2
positive relations with others	3	4	2	5	3
autonomy	2	3	2	5	1
environmental mastery	2	4	1	5	4
purpose in life	4	4	1	5	3
personal growth	3	3	1	5	4

WELLBEING SURVEY AND ASSOCIATION MATRIX TO GET SPECIFIC PARAMETERS (MAX. 3) FROM THE 6 PARAMETERS OF THE RYFF'S SCALE OF PSYCHOLOGICAL WELL BEING

# ADDITIONAL FEATURES (CONTINUED)

## RAG FOR INSIGHTS

- Built a **Retrieval-Augmented Generation** system for mental health wellbeing insights
- Cleaned and combined multiple JSON datasets into a unified instruction-style format
- Converted records into instruction, input, output format for generative querying
- Generated vector embeddings using **SentenceTransformer (all-MiniLM-L6-v2)**
- Indexed embeddings using **FAISS with L2 (Euclidean) distance** for similarity search
- Retrieved top-matching records using FAISS, re-rank them using cross encoder and constructed an augmented prompt based on the context.
- Sent the augmented prompt to **Google Gemini API** to generate wellbeing
- Automatically updated the instruction JSON and FAISS vector store with every new input and Gemini response



OVERVIEW OF RAG FOR  
GENERATING INSIGHTS

# ADDITIONAL FEATURES (CONTINUED)

RUNNING... Stop ⋮

- **Self-Acceptance:** Negative self-image. *Advice: Practice self-compassion; challenge negative thoughts; focus on strengths.*

**Wellbeing Insight Using RAG :**

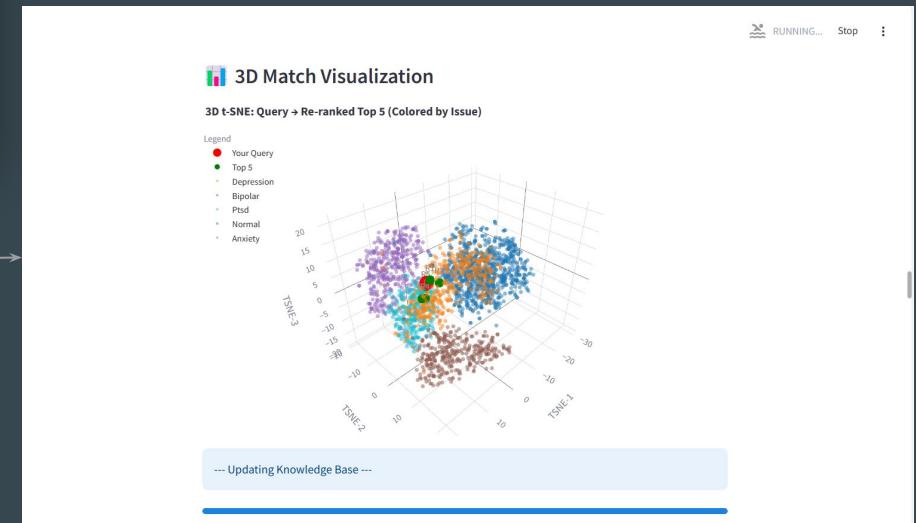
Here's a Ryff Scale wellbeing analysis based on your experiences, tailored to bipolar tendencies:

- **Autonomy:** Mood swings impact sense of control. *Advice: Practice mood charting; identify triggers; implement coping mechanisms.*
- **Environmental Mastery:** Fluctuations hinder daily management. *Advice: Establish a highly structured routine; utilize reminders & task management apps.*
- **Personal Growth:** Cycle fosters negative self-perception. *Advice: Focus on resilience; document coping successes; practice self-forgiveness.*
- **Positive Relations:** Tension and loneliness strain connection. *Advice: Seek support from bipolar-specific communities; transparently communicate needs.*
- **Purpose in Life:** Instability leads to feeling directionless. *Advice: Explore values independent of mood; set small, achievable goals.*
- **Self-Acceptance:** Extreme shifts damage self-image. *Advice: Practice radical self-compassion; challenge mood-related self-criticism.*

**Match Score:**  
0.4896 (higher is better)

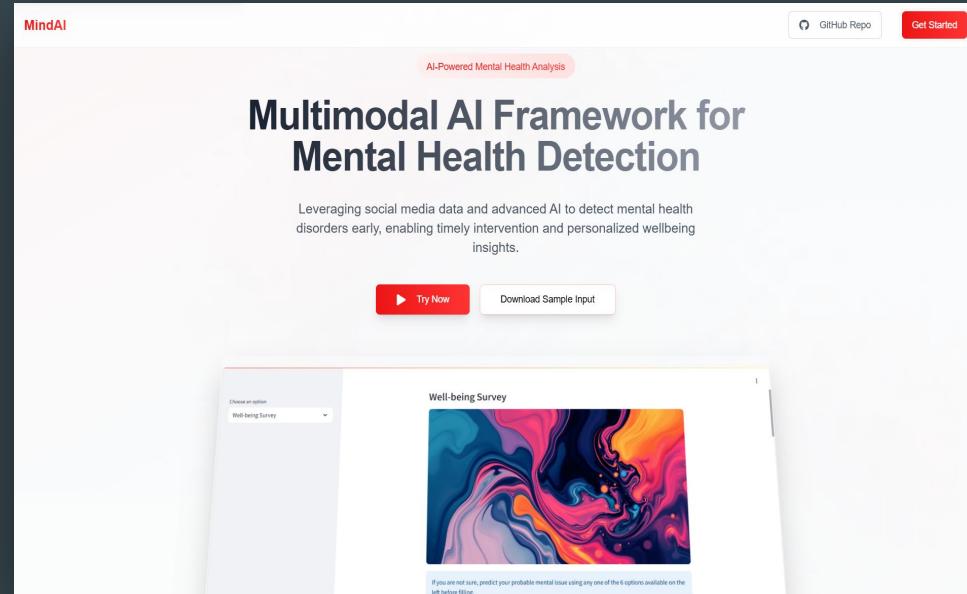
VISUALIZING MATCHES AND KNOWLEDGE BASE

INSIGHTS GENERATED BY RAG



# DEPLOYMENT

- The application requires more resources than the free tier of Streamlit Cloud provides.
- **Main Application:** Runs locally in **Google Colab** for better resource management.
- **Free Static Domain:** Achieved using **ngrok**, allowing external access.
- **Landing Page:** Provides a user-friendly interface to access the application.
- **Sample Inputs:** Available for download via **Google Drive**.



[Landing Page](#)



[Sample Inputs](#)

# 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 and deep learning models, identifying the model with the best performance.

## → *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*

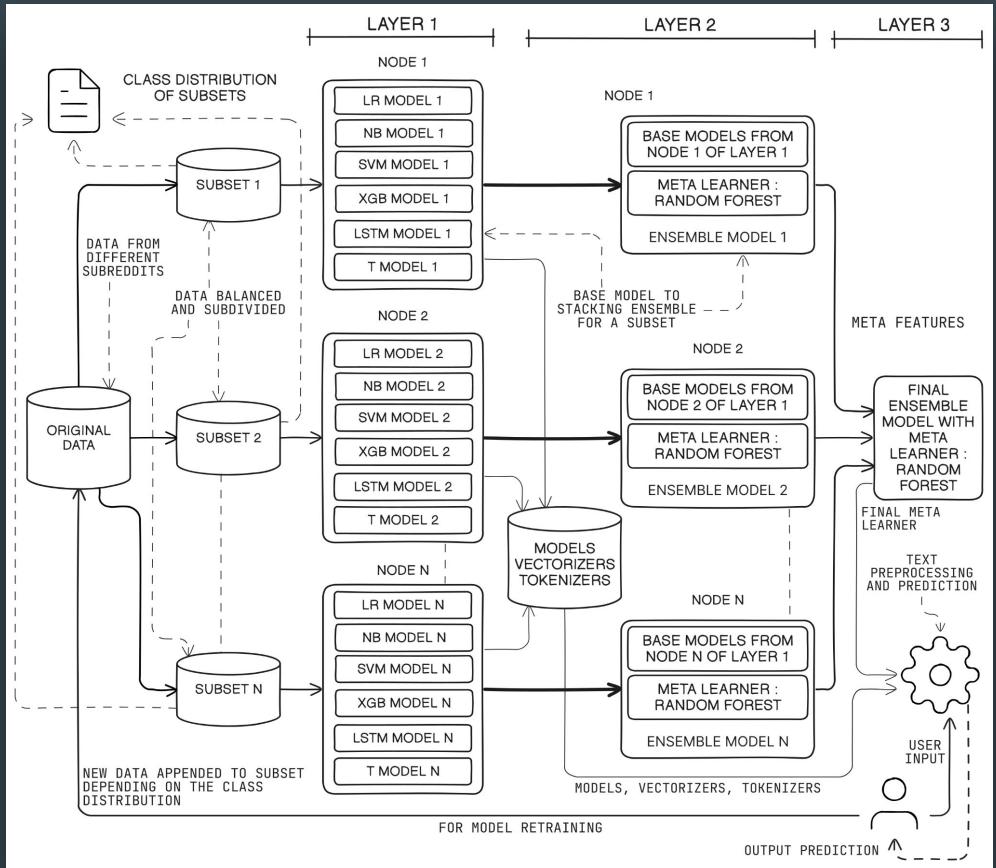
Emphasizes the importance of user privacy and ethical considerations in handling sensitive mental health data.

# FUTURE SCOPE

- Deploy the web application on **cloud platforms** like AWS, Azure, or Google Cloud to enhance scalability and accessibility.
- Implement **caching mechanisms** to store frequently used data or computations, improving response times and computational efficiency.
- Use **threading** to enable simultaneous processing for multiple users, ensuring smooth and efficient application performance.
- Reduce dependency on GPUs by leveraging **distributed systems** for training models across multiple CPUs, improving cost efficiency.
- Optimize the application to **train models on larger datasets** by partitioning the data and distributing workloads effectively across available resources.

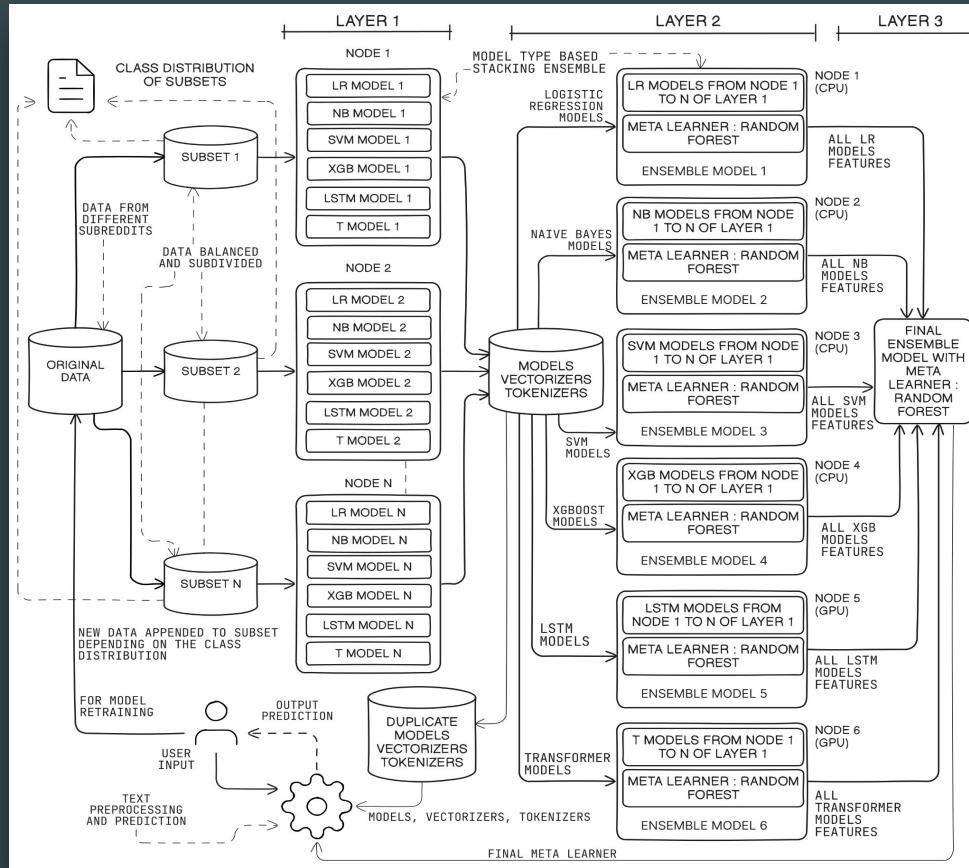


# FUTURE SCOPE (CONTINUED)



ARCHITECTURE 1 , WHERE FINAL ENSEMBLE MODEL IS BASED ON N INTERMEDIATE ENSEMBLE MODELS FROM N SUBSETS

# FUTURE SCOPE (CONTINUED)



ARCHITECTURE 2 , WHERE FINAL ENSEMBLE MODEL IS BASED ON 6 INTERMEDIATE ENSEMBLE MODELS FROM N SUBSETS

# REFERENCES

- **Choudhury, M. D., De, S., & Counts, S. (2013).** Predicting Depression via Social Media. *Proceedings of the 7th International Conference on Weblogs and Social Media*.
- **Guntuku, S. C., Bollen, J., & Lazer, D. (2017).** *Detecting Mental Illness in Social Media*. *American Journal of Public Health*, 107(8), 1279-1285.
- **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.
- **Nadeem, A. (2016).** *A Study of Depression Identification on Twitter*. *International Journal of Computer Applications*, 141(10), 24-28.
- **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.
- **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.
- **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.

# REFERENCES (CONTINUED)

- *Single classifier vs. ensemble machine learning approaches for mental health prediction by Dip Kumar Saha, Tuhin Hossain, Mejdl Safran, Sultan Al Farhood, M. F. Mridha, and Dunren Che* demonstrates that Gradient Boosting (88.80%) and Neural Networks (88.00%) outperform ensemble models (85.60%) for mental health issue prediction, based on survey data from OSMI.
- *Ensemble of hybrid model-based technique for early detection of depression by Konda Vaishnavi, U Nikitha Kamath, B Ashwath Rao, and N V Subba Reddy* introduces a hybrid SVM-MLP model with SMOTE for class balancing, achieving 99.39% accuracy and 99.51% F1-score, highlighting its efficacy in early depression detection.
- *Survey of transformers and towards ensemble learning using transformers for NLP by Hongzhi Zhang and M.Omair Shafiq* reviews transformer models and explores ensemble learning with transformers, showcasing their superior performance in NLP tasks like sentiment analysis and text generation.



THANK YOU