MENTAL HEALTH ISSUE IDENTIFICATION

Leveraging Technology for Early Detection and Support

Team Members

- Isaac Wanganga
- Beryl Wafula
- Jim Akoko
- Victor Maina
- Bestina Nyiva
- Cynthia Jerono

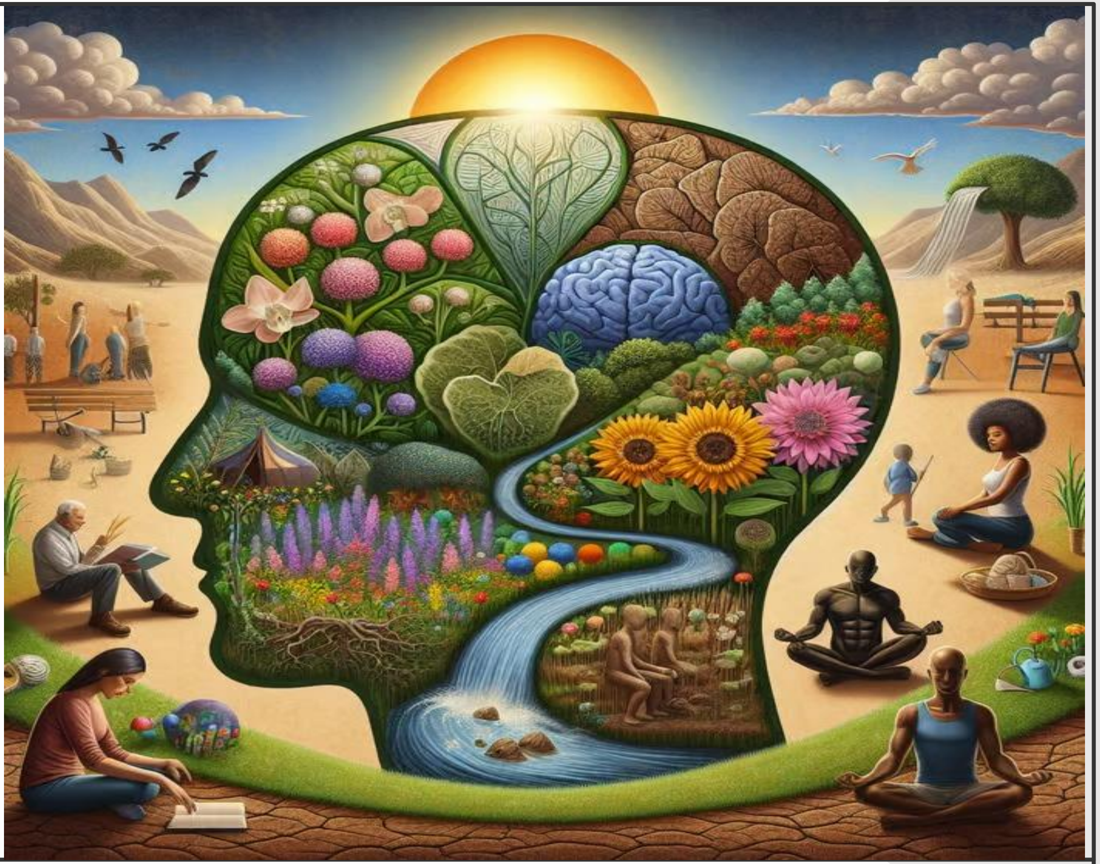


Table of contents

01

Business Understanding

Problem statement

Mental health issues are rising globally, yet many individuals lack access to timely diagnosis and support.

This project aims to develop a system that identifies mental health concerns using data-driven methods, enabling early intervention and improved outcomes.

Goals and Objectives

Early Detection of Mental Health Issues

Develop a system to identify signs of mental health concerns at an early stage using predictive analytics and data-driven insights.

Improve Accessibility to Support

Create a platform that bridges the gap between individuals and mental health resources, promoting timely intervention.

Raise Awareness and Reduce Stigma

Use the system to educate users on mental health issues, fostering a culture of understanding and acceptance.

Stakeholders

Healthcare Providers

Mental health professionals who will utilize the system for early diagnosis and treatment planning.

Individuals at Risk

People experiencing mental health challenges who benefit from timely identification and intervention.

Policy Makers and Organizations

Authorities and institutions focused on improving mental health services and awareness in communities.

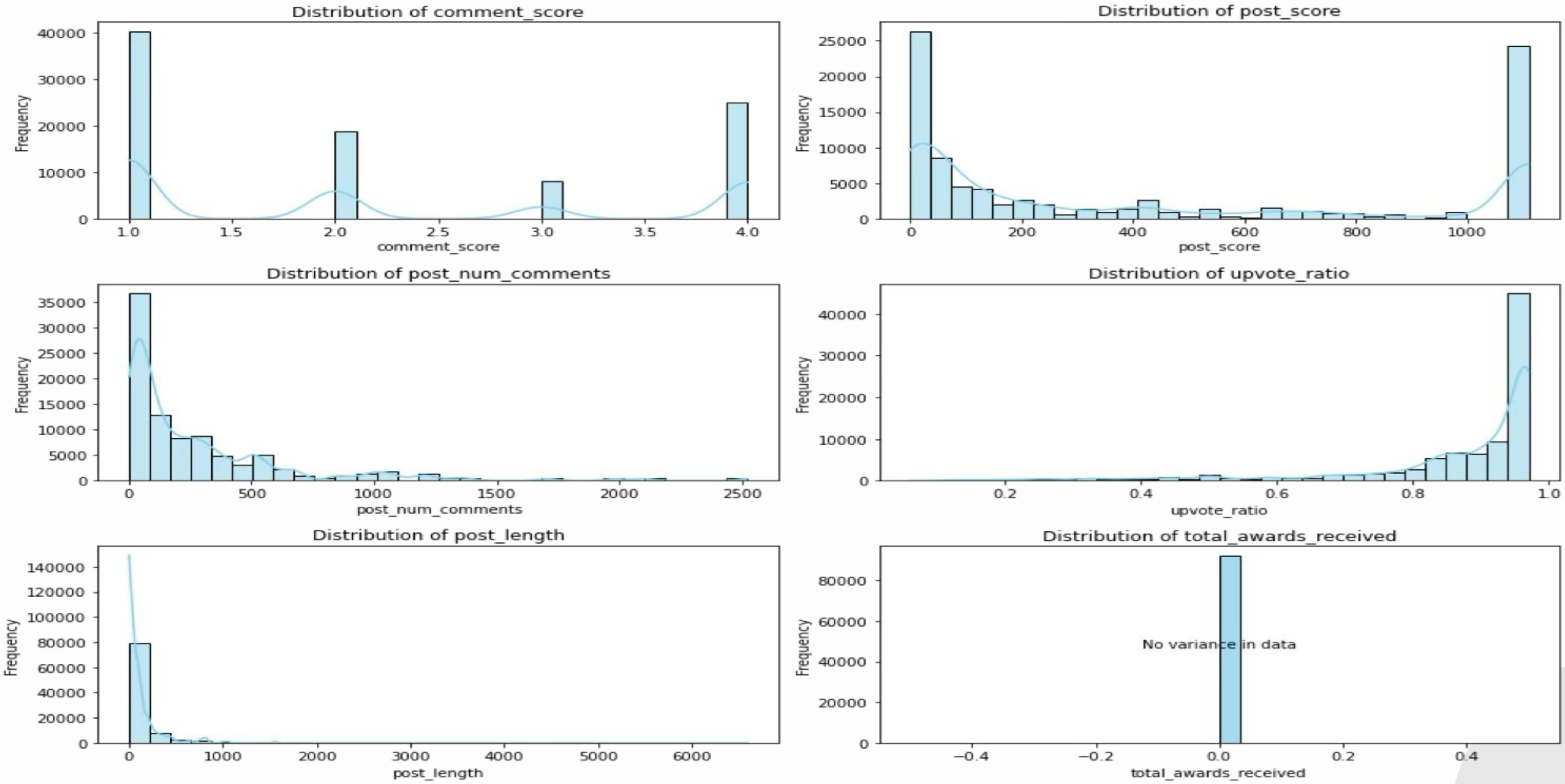
Libraries used

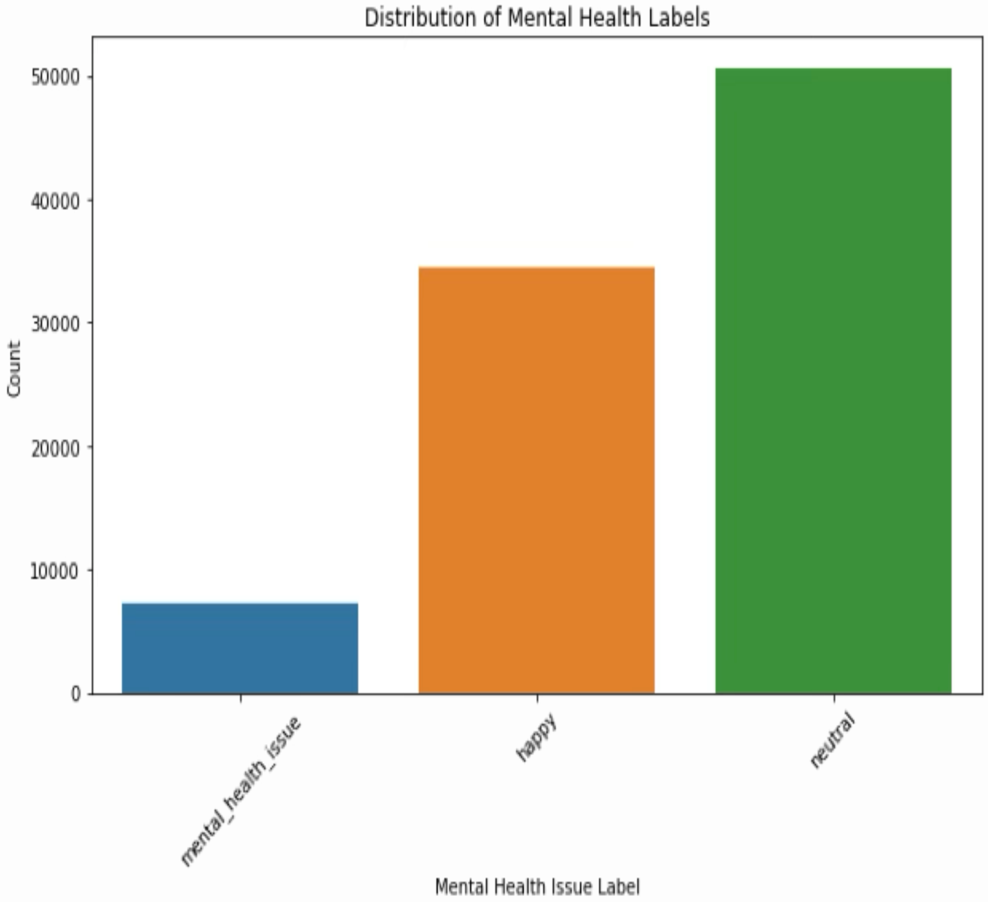
02

Data Understanding

- **Source: Data was collected using the python reddit API Wrapper (PRAW) .**
- **Methodology: Keyword-based search queries were used to scrape up to 5,000 posts per subreddit related to mental health, positive expressions, and neutral content.**
- **Output: The data was structured and saved as a CSV file, enabling comprehensive analysis of mental health discussions, sentiment, and engagement patterns.**
- **Attributes Captured: Post titles, bodies, comments, metadata (e.g., author info, comment scores, timestamps, flair, upvote ratios, and crosspost counts).**

Distribution of numerical features





The chart shows three categories with varying frequencies. Neutral category (the green bar) has the highest count, followed by the happy/positive category (orange bar), and the lowest is the mental health issue category (blue bar).

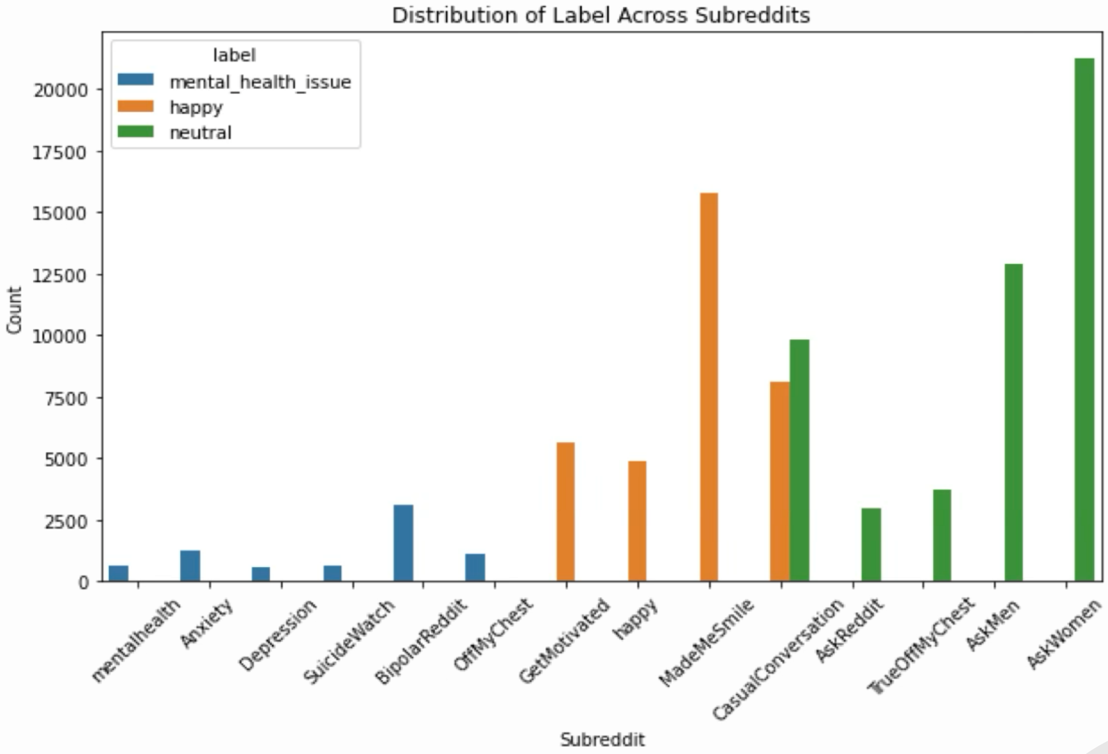
Implications

some categories are more prevalent in the dataset e.g Neutral and happy . This could impact model training, as an imbalanced dataset may lead the model to perform better on the majority category and worse on the minority.

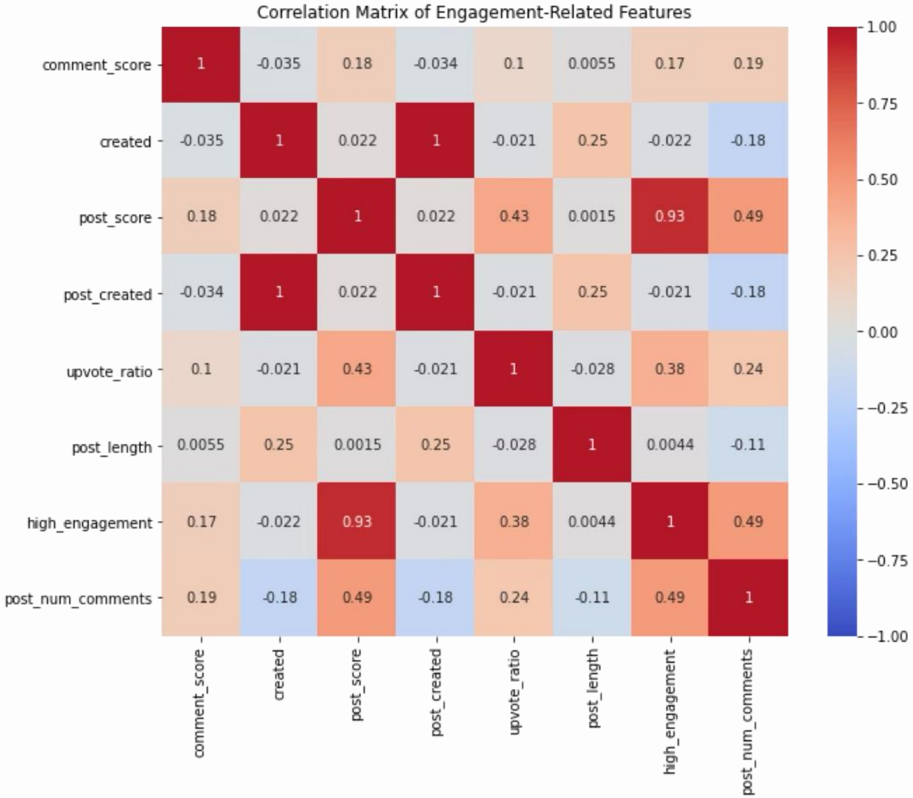
Insights:

Subreddits with a higher proportion of "happy" or "neutral" labels might reflect communities with generally positive or balanced sentiment.

In contrast, subreddits with more "mental_health_issue" posts could represent spaces where users share struggles, indicating their value for mental health support discussions.



Correlation Matrix for Numerical Features



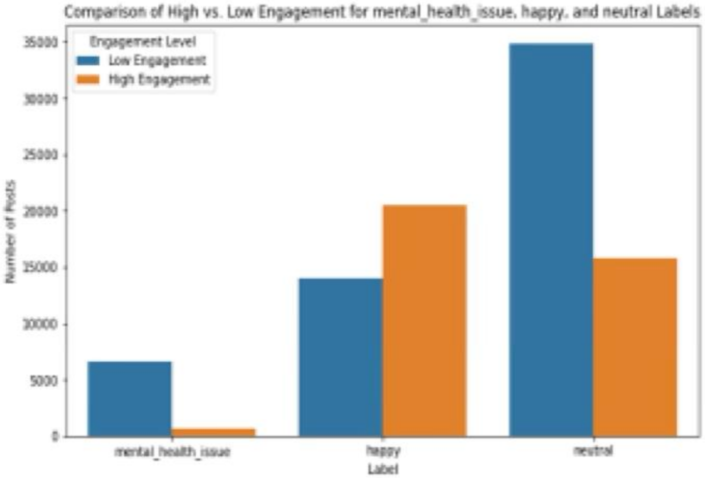
Most numerical features have low collinearity ranging from -0-0044 to 0.49.
The features with high multicollinearity are post created and created with a collinearity of 1

Distribution of high and Low Engagement per label

1.Neutral content has the highest number of low-engagement posts (blue), indicating that neutral topics generally receive less interaction from the audience.

2.Mental health-related content shows low numbers of hign engagement posts indicating that people with mental health issues most likely don't want to talk about it on social platform hence the high number posts with low engagement .

3.The Happy label has both moderate number of both high and low engagement posts, indicating a balanced engagement level for positive content.



Popular Words Per Label

Mental Health Issue

"know, feel, want, life, mental, health"

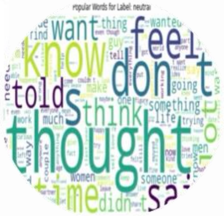
Happy

"know, love, make, feel, want, work"

Neutral

"thought, feel, know, said, want, time"

The words like "know," "feel," and "want" are common across labels, indicating that discussions may revolve around introspection, personal desires, and emotions.



03

Data Preparation

Steps taken to prepare the data

Data Loading

Imported datasets using Python libraries (e.g., pandas).

Data Cleaning

Handled missing values by imputation or removal and remove duplicates for consistency.

Feature Engineering

Created new features and transformed existing ones as needed. Also encoded categorical values

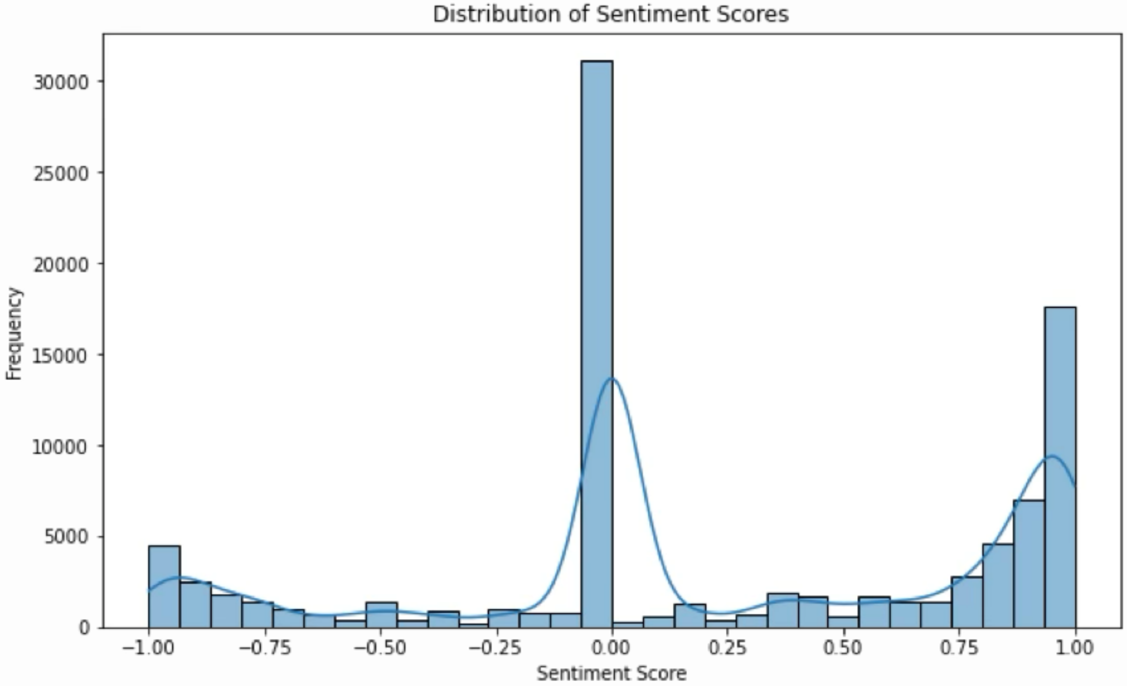
Data Transformation

Encoded categorical variables and scaled numerical data for machine learning compatibility.

Splitting the Data

Split the data into training and testing sets.

Feature Engineering



Sentiment Score

The first peak represents one group whose sentiment score is centered around -0.25 (fairly negative) while the second peak represents a different group whose sentiment score is centered around 1(very positive).

Implications

- Given the clear separation between the two peaks, it may be beneficial to treat the two groups separately in our analysis.
- This segmentation could allow for more targeted insights or better model performance, especially if the behaviors or language used in each group differ.

04

Modelling

Models used

- **Models**
  - Logistic Regression
  - Random Forest Classifier
  - Decision Tree
  - DistilBERT

Models review

- Logistic Regression
  - **Accuracy**: Moderate, but high misclassification for detecting "mental health issues."
  - **Insights**: Poor identification of critical cases due to many false negatives.
- Random Forest Classifier
  - **Accuracy**: Improved compared to Logistic Regression.
  - **Performance**: More balanced with fewer false negatives for "mental health issues."
- Decision Tree
  - **Performance**: Nearly perfect, with minimal misclassifications.
  - **Concerns**: High accuracy suggests possible overfitting.
- DistilBERT (Transformer-based model for text classification)
  - **Accuracy**: **90.8%** on the evaluation set.
  - **Insights**: Strong generalization and effective for detecting mental health-related issues.

Model Scores

Models	Accuracy	Precision	Recall	F1-score
Logistic Regression	77.5%	74%	77%	76%
Random Forest Classifier	95.3%	97%	92%	95%
Decision Tree	98.8%	94%	96%	95%
DistilBERT	98.8%	99%	99%	99%

Model Performance Comparison

The logistic regression has an overall accuracy of 81%. However, it struggles with identifying class 1 (mental health issues).

Random Forest performs exceptionally well, achieving 100% precision and 78% recall for class 1(mental Health Issue). This means it perfectly identifies all mental health issues without misclassification.

Decision tree has an accuracy score of 0.989 for a decision tree model indicates that the model is highly successful at predicting the correct labels in your dataset. However the could be chances of overfitting.

Distilbret model

```
Logistic Regression Accuracy: 0.7752200894790013
Logistic Regression Classification Report:
               precision    recall  f1-score   support

           0       0.74      0.77      0.76     10350
           1       0.63      0.12      0.20      2173
           2       0.80      0.87      0.84     15193

    accuracy                           0.78     27716
   macro avg       0.72      0.59      0.60     27716
weighted avg       0.77      0.78      0.76     27716

Random Forest Accuracy: 0.9532400057728387
Random Forest Classification Report:
               precision    recall  f1-score   support

           0       0.97      0.92      0.95     10350
           1       1.00      0.78      0.87      2173
           2       0.94      1.00      0.97     15193

    accuracy                           0.95     27716
   macro avg       0.97      0.90      0.93     27716
weighted avg       0.96      0.95      0.95     27716

Decision Tree Accuracy: 0.988995526049935
...
    accuracy                           0.99     27716
   macro avg       0.99      0.99      0.99     27716
weighted avg       0.99      0.99      0.99     27716
```

DISTILBERT Evaluation

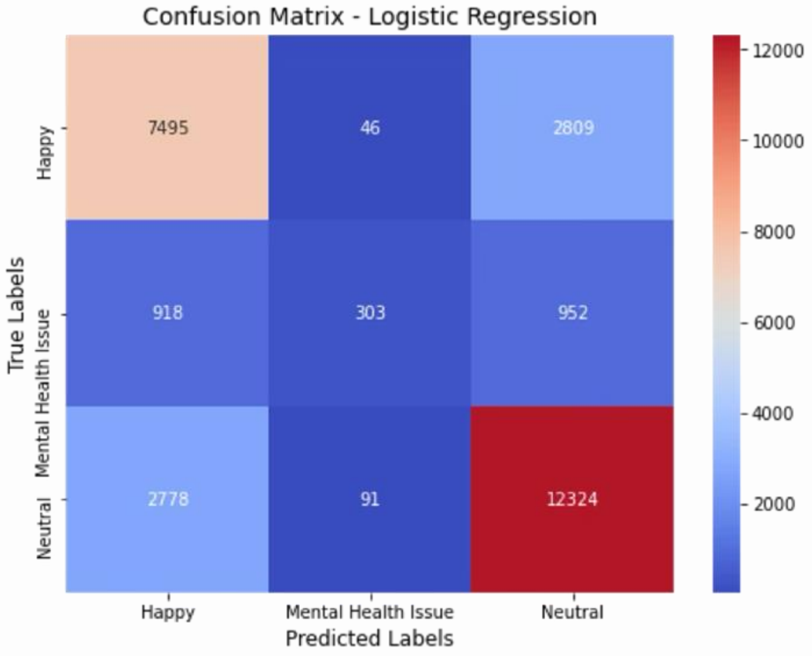
Evaluation Metrics:
Accuracy: Achieved 90.8% on the evaluation set, reflecting strong model performance.
Loss: Final evaluation loss was 0.33, indicative of a good fit but with room for further optimization.

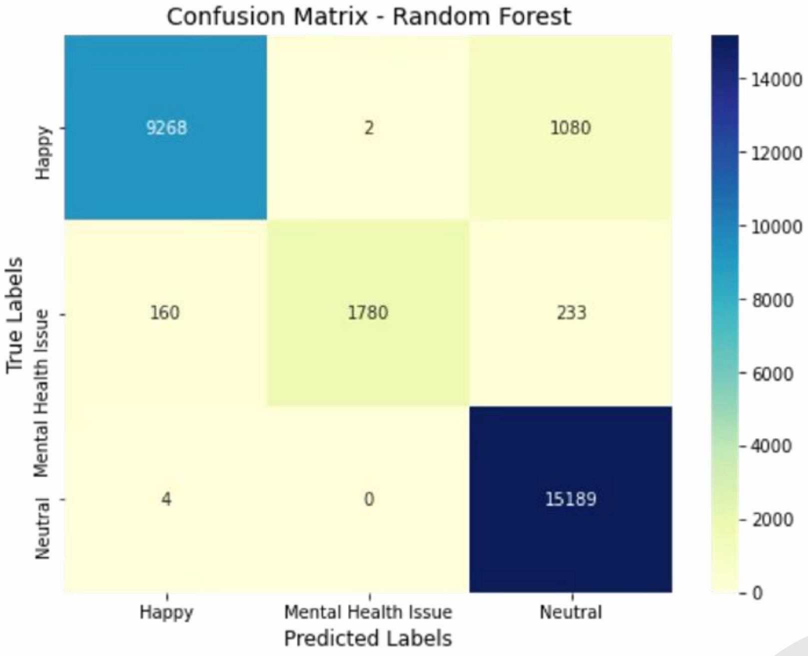
Classification Report:
Class-level Performance:
Happy: Precision (90%), Recall (83%), F1-score (87%).
Mental Health Issue: Precision (89%), Recall (94%), F1-score (91%).
Neutral: Precision (89%), Recall (93%), F1-score (91%).

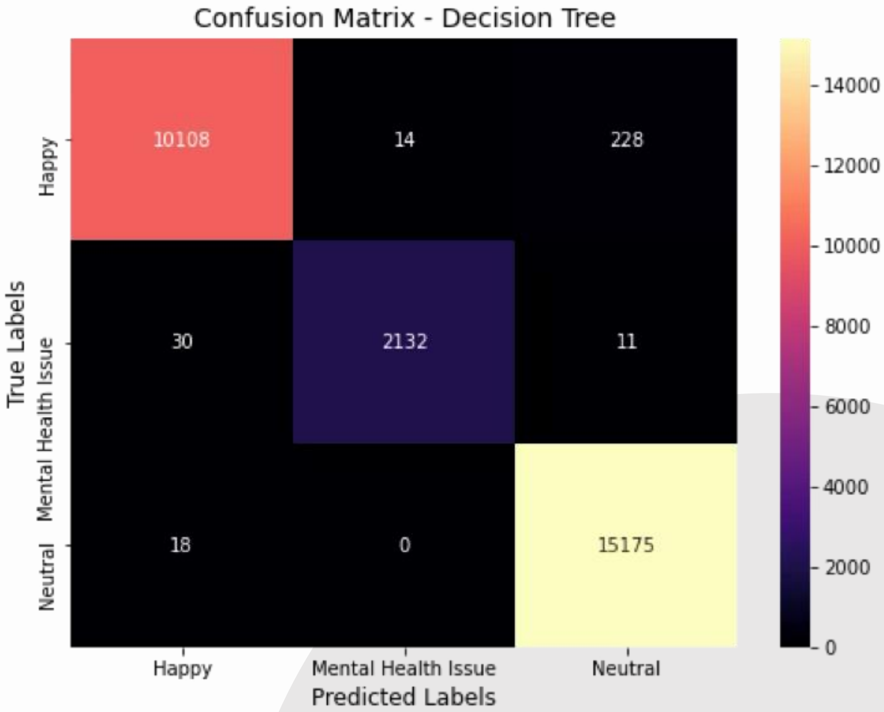
Class imbalances appear well-managed, with robust performance across all categories.
High recall for detecting mental health issues (94%) ensures fewer critical cases are missed.
Balanced performance across all classes, with good precision and recall synergy.

```
100%|██████████| 3639/3639 [14:59:21<00:00, 14.83s/it]
{'train_runtime': 53961.6592, 'train_samples_per_second': 1.079, 'train_steps_per
100%|██████████| 520/520 [10:46<00:00,  1.24s/it]
Evaluation results: {'eval_loss': 0.2611389458179474, 'eval_accuracy': 0.89392663
100%|██████████| 520/520 [11:08<00:00,  1.29s/it]
DistilBERT Classification Report:
                   precision    recall  f1-score   support

            happy       0.90      0.83      0.87      3152
mental_health_issue       0.89      0.94      0.91       673
          neutral       0.89      0.93      0.91      4490

         accuracy                           0.89      8315
        macro avg       0.89      0.90      0.90      8315
     weighted avg       0.89      0.89      0.89      8315
```

Confusion Matrices







Logistic Regression: The model showed High misclassification for class 1 (mental health issues) with many false negatives(1871), indicating poor identification of this class.

Random Forest: The model showed Better performance for class 1 (mental health issues) , though few misclassifications still (393 False negatives). Overall, more balanced results compared to Logistic Regression.

Decision Tree: The Model Showed Nearly perfect performance for all classes, with minimal misclassifications, making it the most effective model for identifying mental health issues. However, this near perfect classification is a sign of overfitting

Model chosen

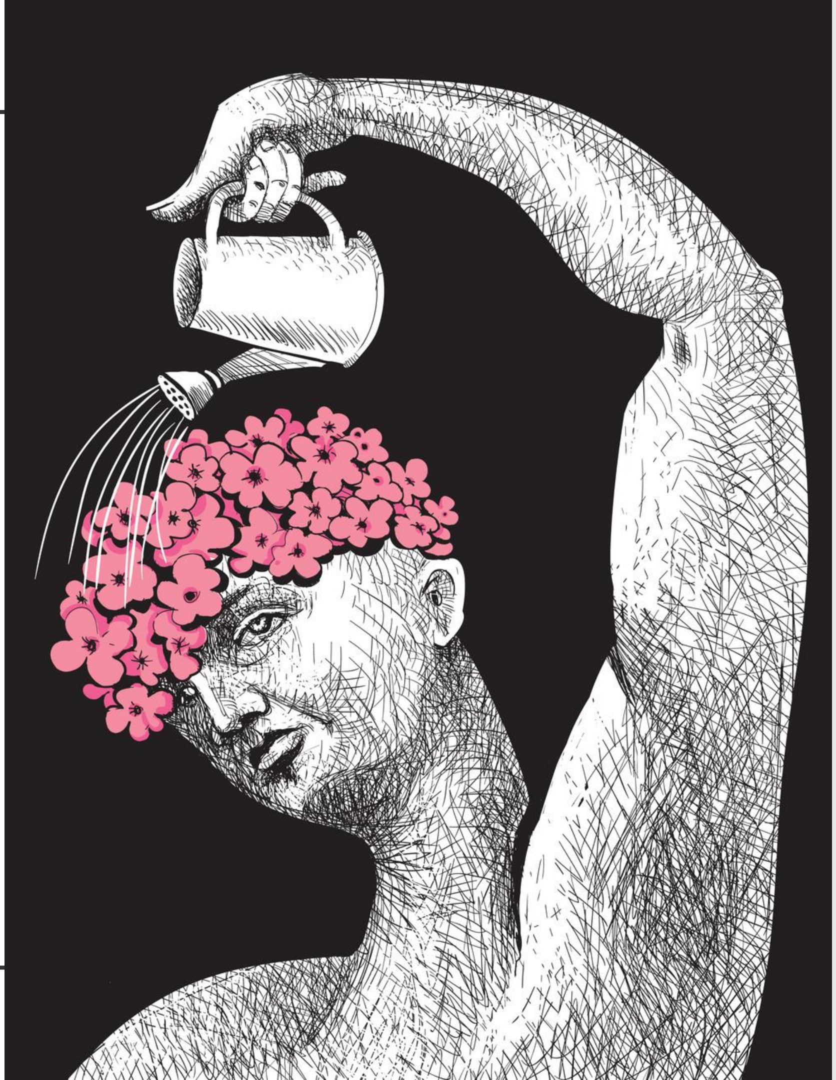
DistilBERT was chosen for deployment due to its robust handling of text data and superior performance metrics across all classes.

Limitations of the project

- Limited Generalizability: Social media users may not represent the broader population, especially for mental health issues. For instance, younger individuals or those more comfortable with online platforms may dominate the dataset.
- Subreddit-Specific Norms: Subreddits have unique cultures and posting behaviors, which may bias the interpretation of "mental_health_issue," "happy," or "neutral" labels.
- Self-Selection Bias: Users discussing mental health issues on social media are often those already seeking help or validation, which may skew the dataset toward specific types of mental health issues or perspectives.

06

Deployment



Steps

01

Frontend

The model was deployed on streamlit and a link was generated for users to view the model.

02

Model

The model was saved in a pickle file to enable integration with the backend.

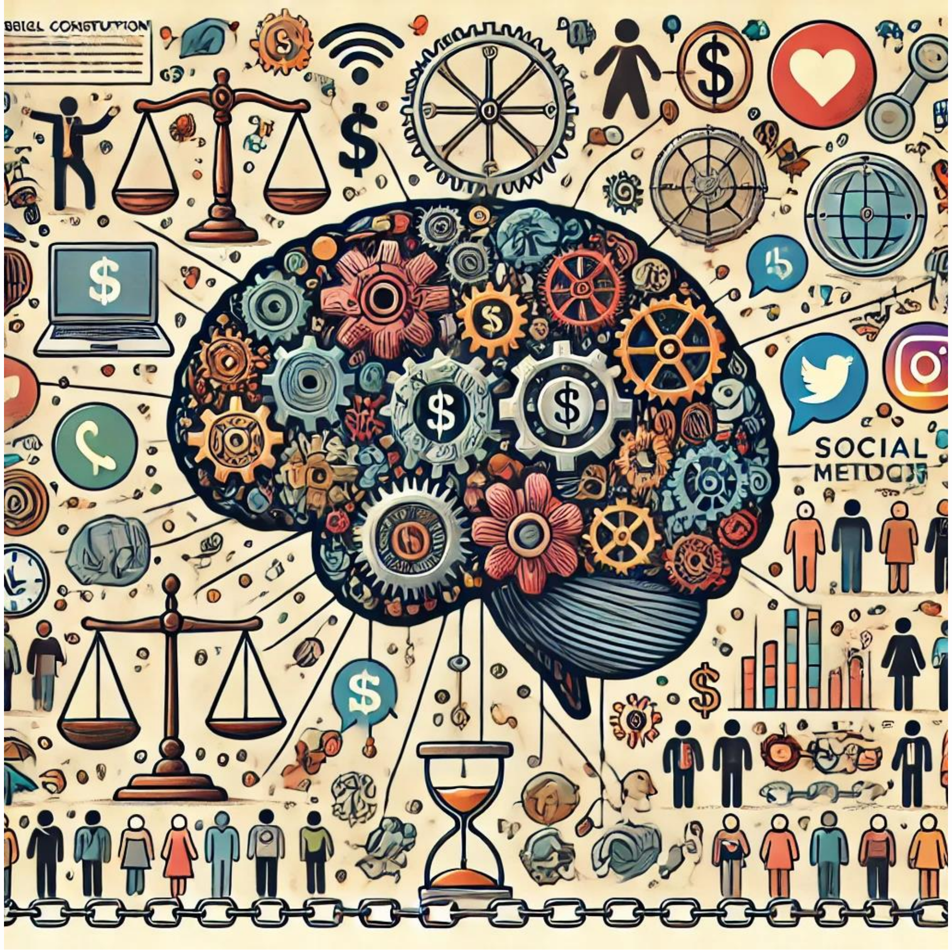
03

Backend

Fast api was used to create the backend and was integrated with the model to get responses.

07

Conclusion



- Our project demonstrates the potential of AI in addressing critical mental health challenges by leveraging advanced models like DISTILBERT to analyze social media text data.

- The integration of WebSocket technology also ensures seamless user experience, while the focus on intervention strategies underscores the ethical responsibility of using AI for societal impact.

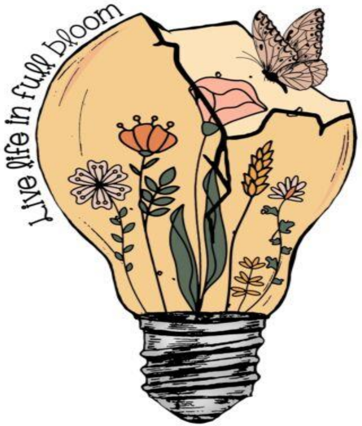
- Our project highlights the importance of continuous refinement, maintaining ethical considerations such as data privacy , and collaboration with mental health professionals to create a meaningful and impactful solution.

08

Recommendations

1.Model Optimization

We will fine-tune our DISTILBERT model by perform hyperparamet optimization (e.g., learning rate, epochs, batch size) to improve its performance across all classes.



2.Comprehensive Reporting for Mental Health Monitoring and Intervention Strategies

Generate detailed reports or summaries (e.g., user-specific insights, community-level trends) that can be shared with mental health professionals or moderators. Include potential intervention strategies based on high-risk categories.

3.Integrate External Mental Health Resources

Collaborate with mental health advocacy groups and policymakers to utilize insights for targeted awareness campaigns and community-level interventions.

Thanks!

Do you have any questions?