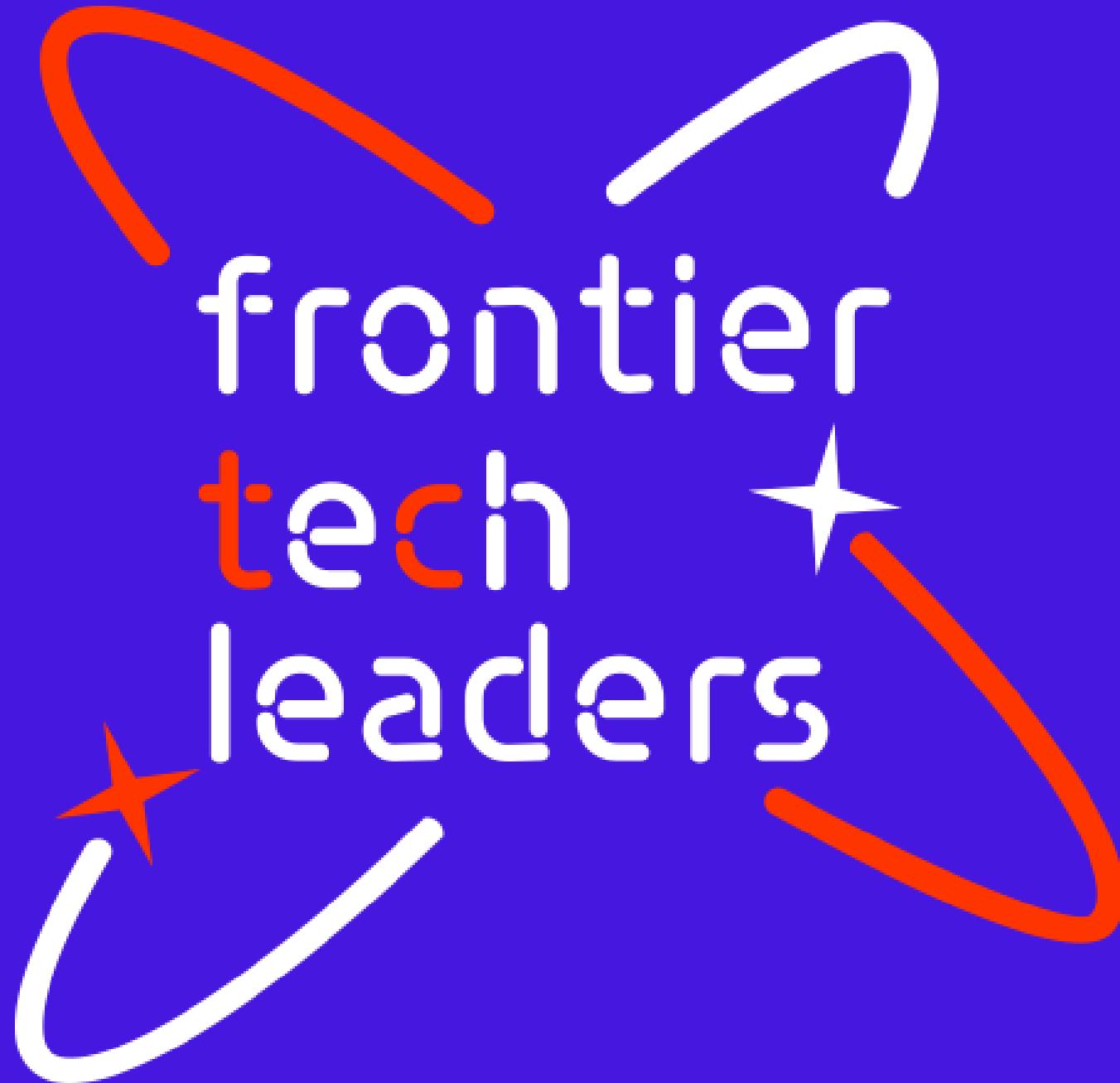


A new generation  
of tech **specialists**





# Mental Health Detection from Social Media Posts

By Group 12



# Background

- **Context:** Mental health is a critical component of global well-being. Platforms like Facebook and X are primary forums for personal expression where users share emotions, creating a massive data stream where distress signals are often lost.
- **Problem:** Traditional manual support and reactive reporting tools cannot handle the volume of millions of daily posts. Subtle linguistic patterns of distress often go unnoticed until a crisis becomes apparent.
- **Solution Necessity:** An automated NLP-driven approach is required to analyze text at scale to bridge the gap between expression of distress and professional support.

# Objectives

- The machine learning model aims to automatically detect and classify mental health conditions from social media text by analyzing users' posts in real time.
- Using a fine-tuned DistilBERT model, it processes textual data to identify patterns related to mental health states (such as anxiety or normal behavior) and provides fast, reliable predictions through a deployed cloud API.

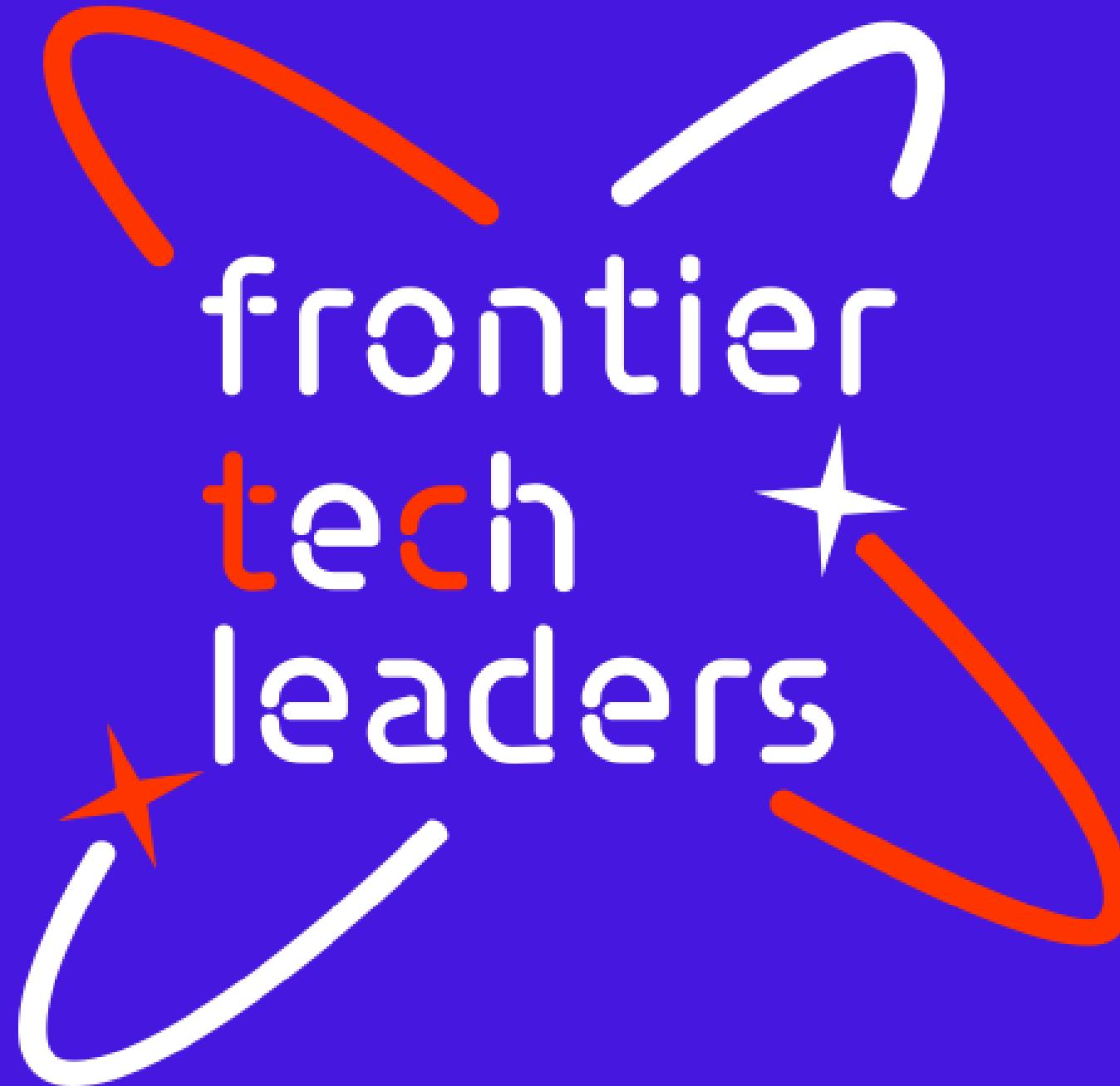
# SDG Relation

## SDG 3: Good Health and Well-Being

By supporting early awareness of mental-health risk patterns, the system aims to assist organizations in promoting mental well-being and prevention, while avoiding harmful automation.

## SDG 9 – Industry, Innovation, and Infrastructure

By leveraging modern natural language processing models and cloud-based APIs, the system demonstrates how AI-driven infrastructure can be applied responsibly to analyze large-scale social-media data. The project emphasizes scalable, modular design and supports innovation in digital monitoring tools while maintaining human oversight in sensitive decision-making.



Data

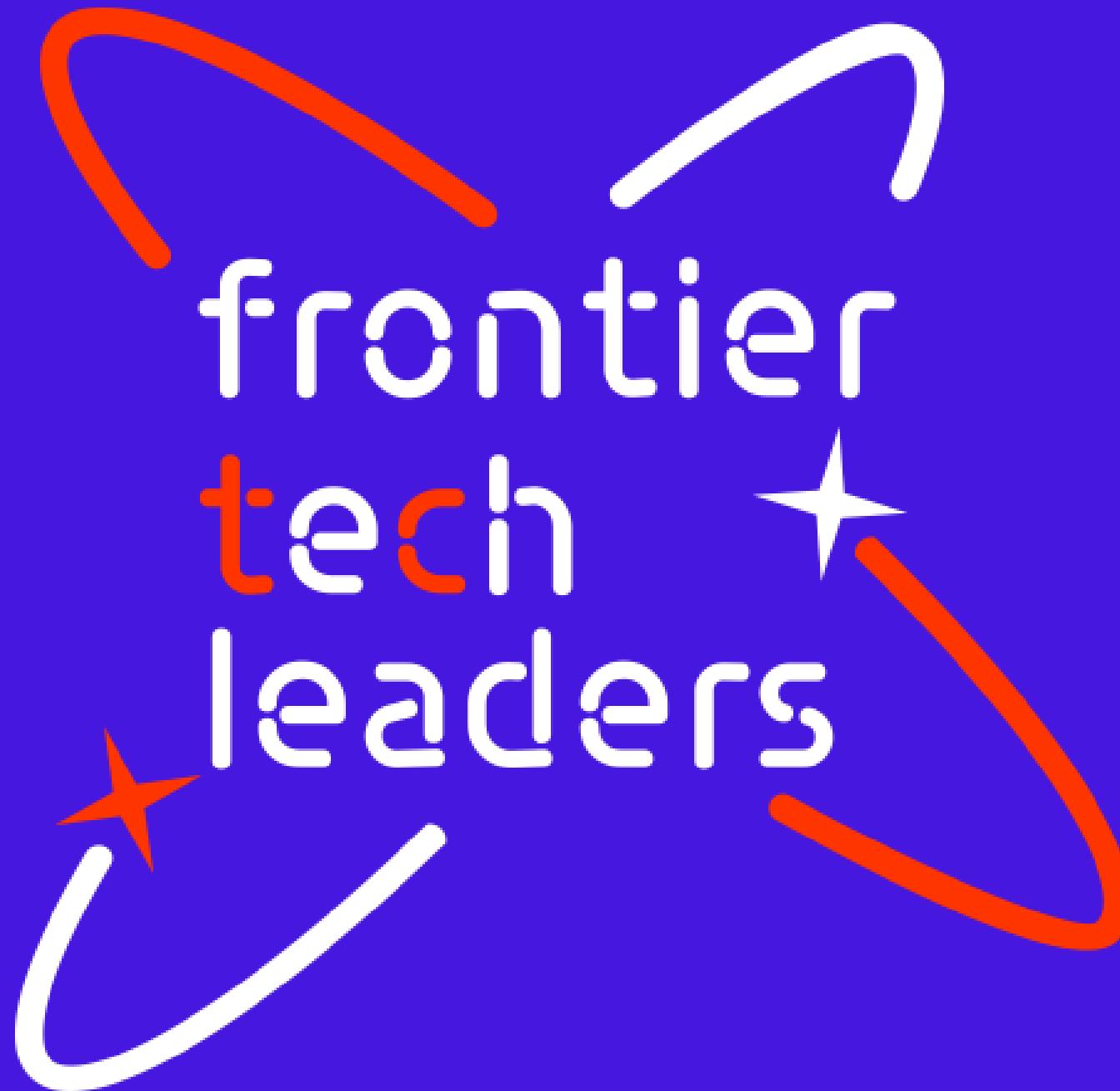


# Data

- **Primary Dataset**
  - Source: Kaggle – Mental Health Sentiment Analysis
  - Link: <https://www.kaggle.com/code/annastasy/mental-health-sentiment-analysis-nlp-ml/input>
- Live social-media posts (scraped via API for deployment testing)
- **Data Characteristics**
  - Short, informal text
  - Emotionally expressive language
  - Multiple mental-health categories

# Data Preprocessing

- Text cleaning and normalization
- Removal of empty or invalid posts
- Translation of non-English (e.g., Burmese) text into English for model compatibility
- Preservation of original text for human review and transparency



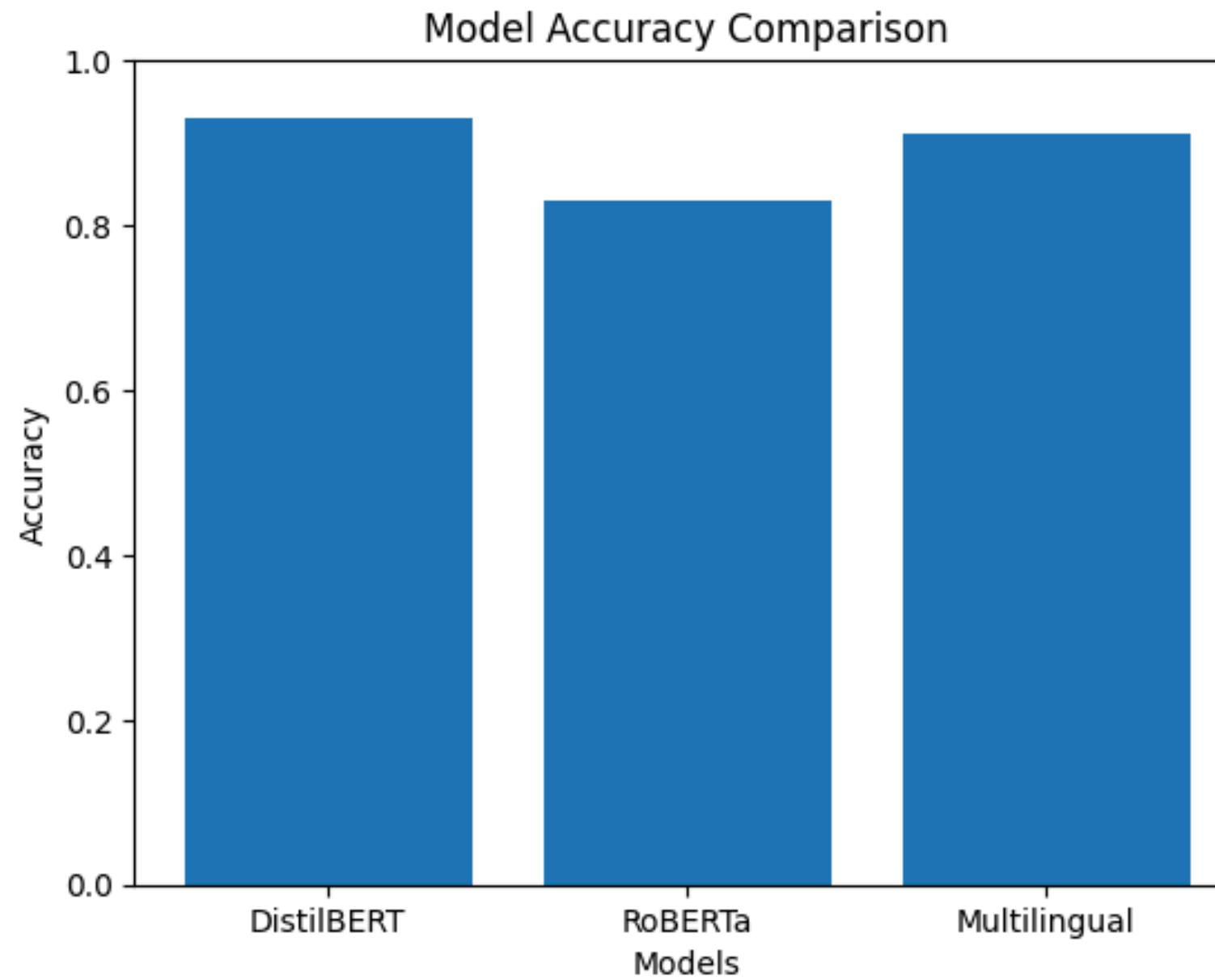
Model



# Model Selection

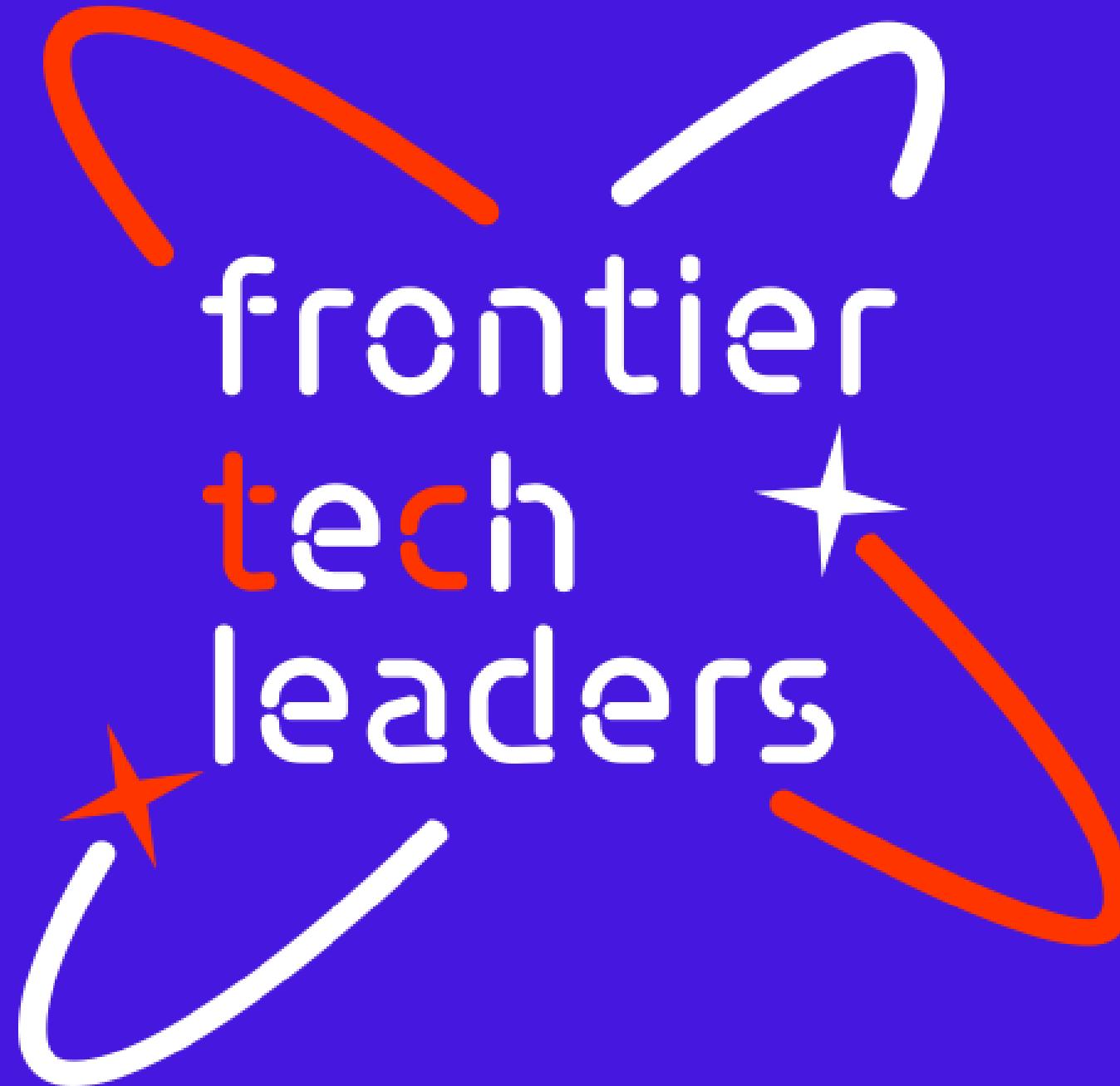
- **Baseline Approach:**
  - Initially experimented with classic models (Logistic Regression, XGBoost) using TF-IDF features .
  - Limitation: Failed to capture context (e.g., distinguishing "not happy" from "happy") due to lack of sequence awareness .
- Selected Architecture: **DistilBERT (distilbert-base-uncased)**
  - Contextual Intelligence: Uses self-attention to understand the relationship between words across the entire sentence.
  - Efficiency: Chosen for its superior balance of high accuracy (~93%) and faster inference speed compared to larger models like RoBERTa.

# Models Comparison



# Model Training and testing

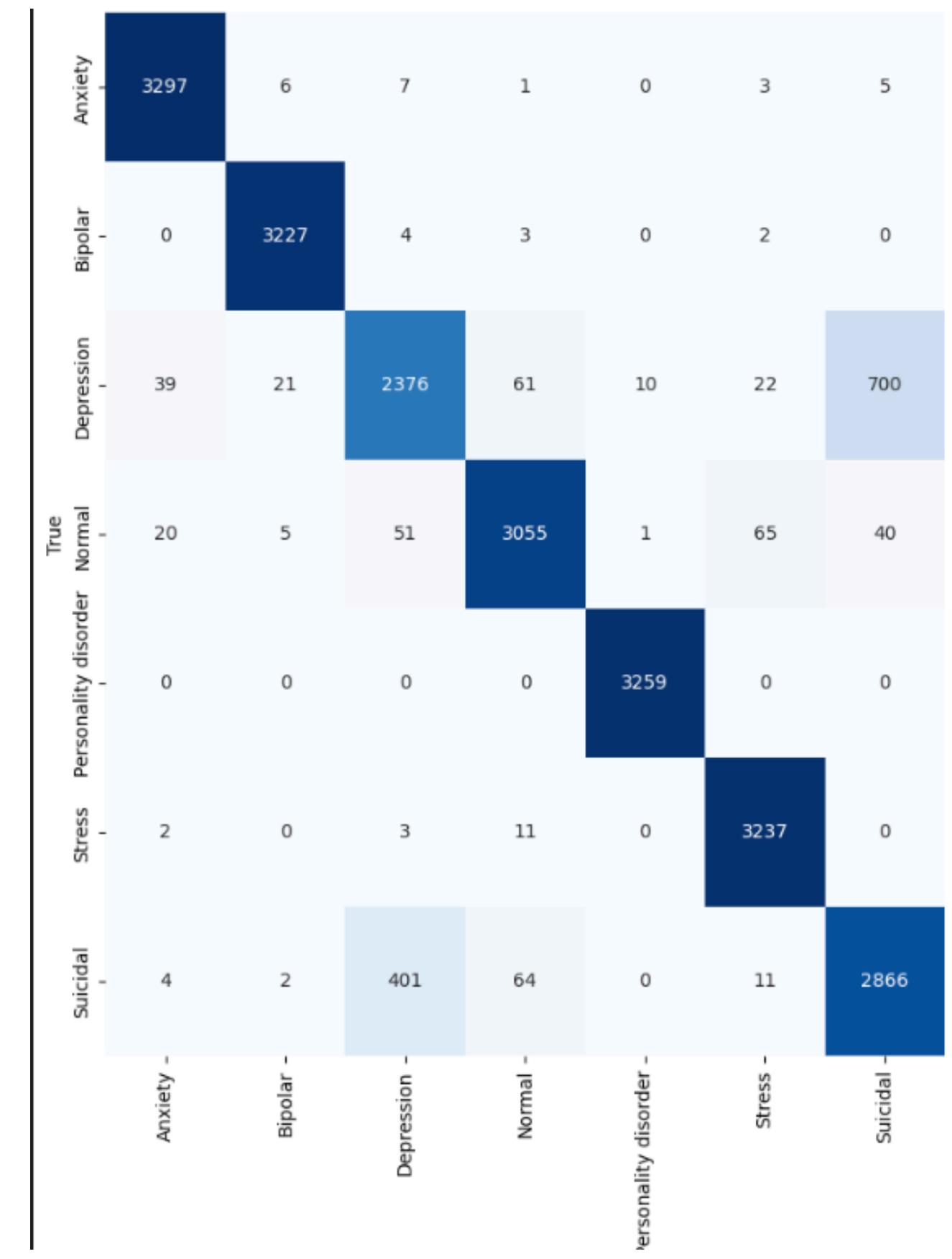
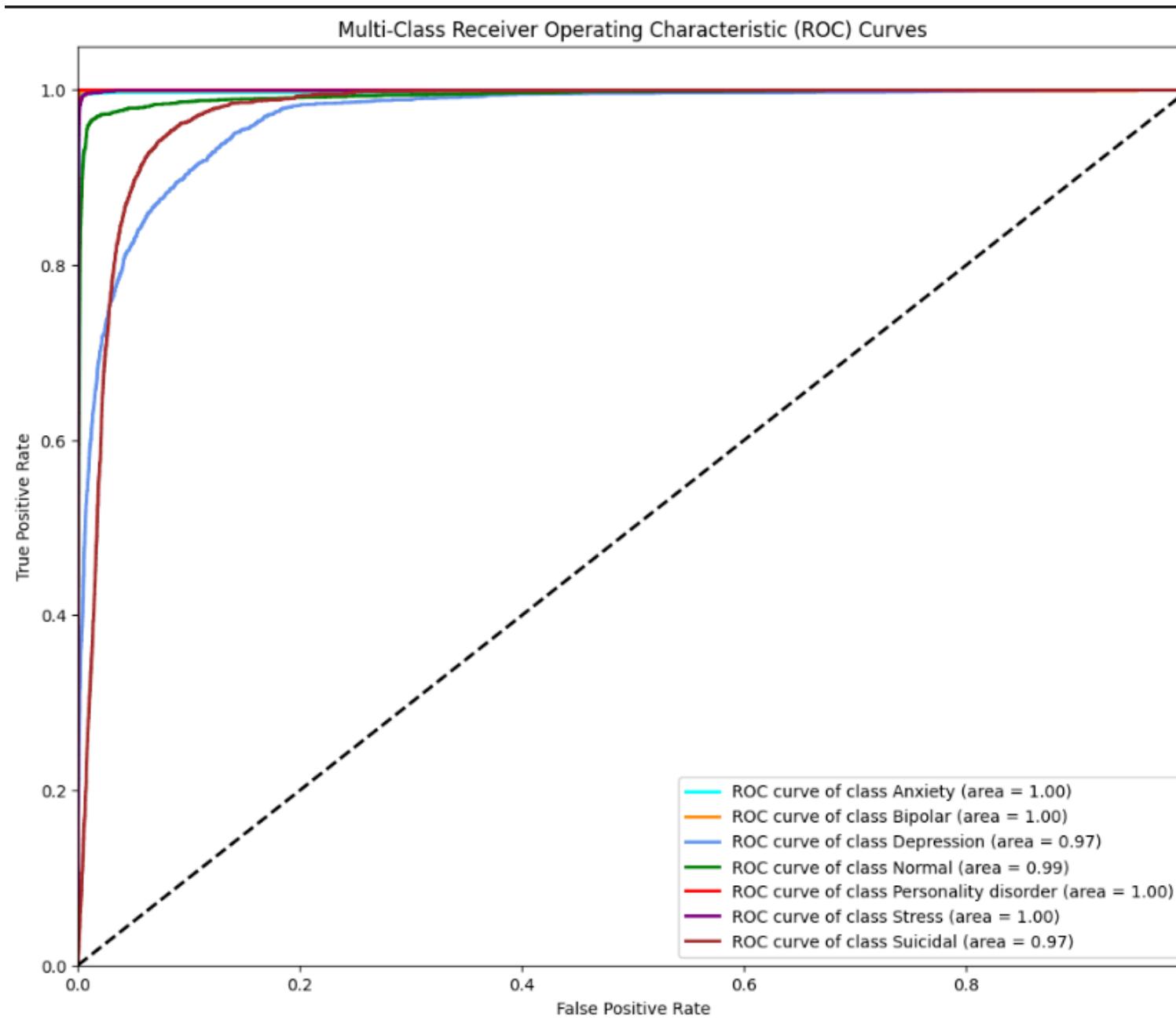
- **Training Configuration:**
  - **Tokenization:** Switched to WordPiece to break words into meaningful sub-words, preserving semantic value better than stemming .
  - **Optimization:** Utilized Optuna to verify optimal learning rates  $1 \times 10^{-5}$  to  $5 \times 10^{-5}$
  - **Execution:** Fine-tuned for 2 epochs with a batch size of 16 using the Hugging Face Trainer API.
- **Testing Protocol:**
  - **Data Split:** Evaluated on a 20% held-out test set strictly unseen during training to ensure valid performance metrics.
  - **Primary Metric:** Prioritized Weighted F1-Score to ensure a safe balance between Precision (accuracy of flags) and Recall (catching true cases).



# Result

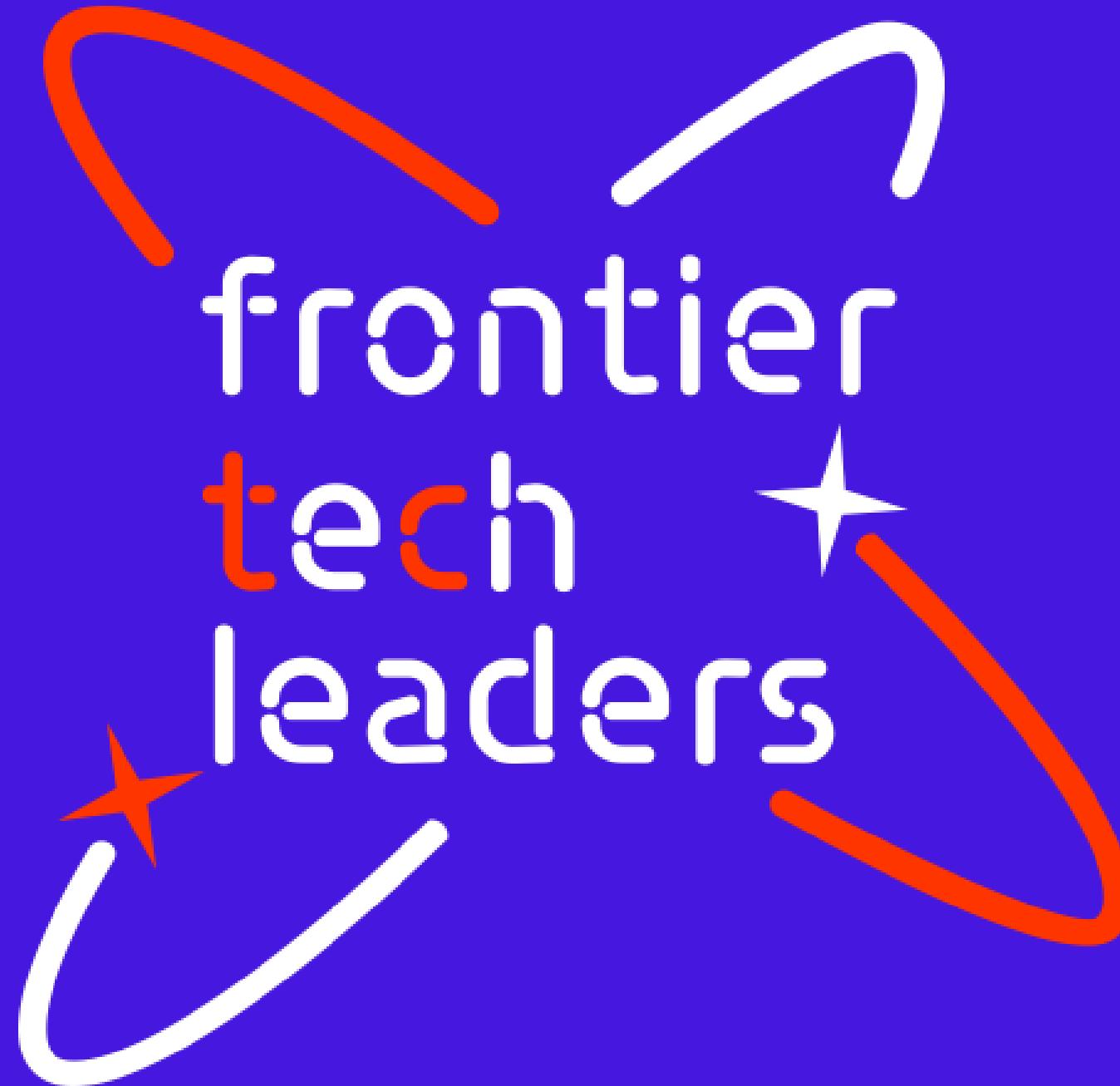


# Evaluation Results



# Classification Report

	Precision	Recall	F1 -Support
Anxiety	0.98	0.99	0.99
Bipolar	0.99	1.00	0.99
Depression	0.84	0.74	0.78
Normal	0.96	0.94	0.95
Personality Disorder	1.00	1.00	1.00
Stress	0.97	1.00	0.98
Sucidial	0.79	0.86	0.82
Test Accuracy			0.93



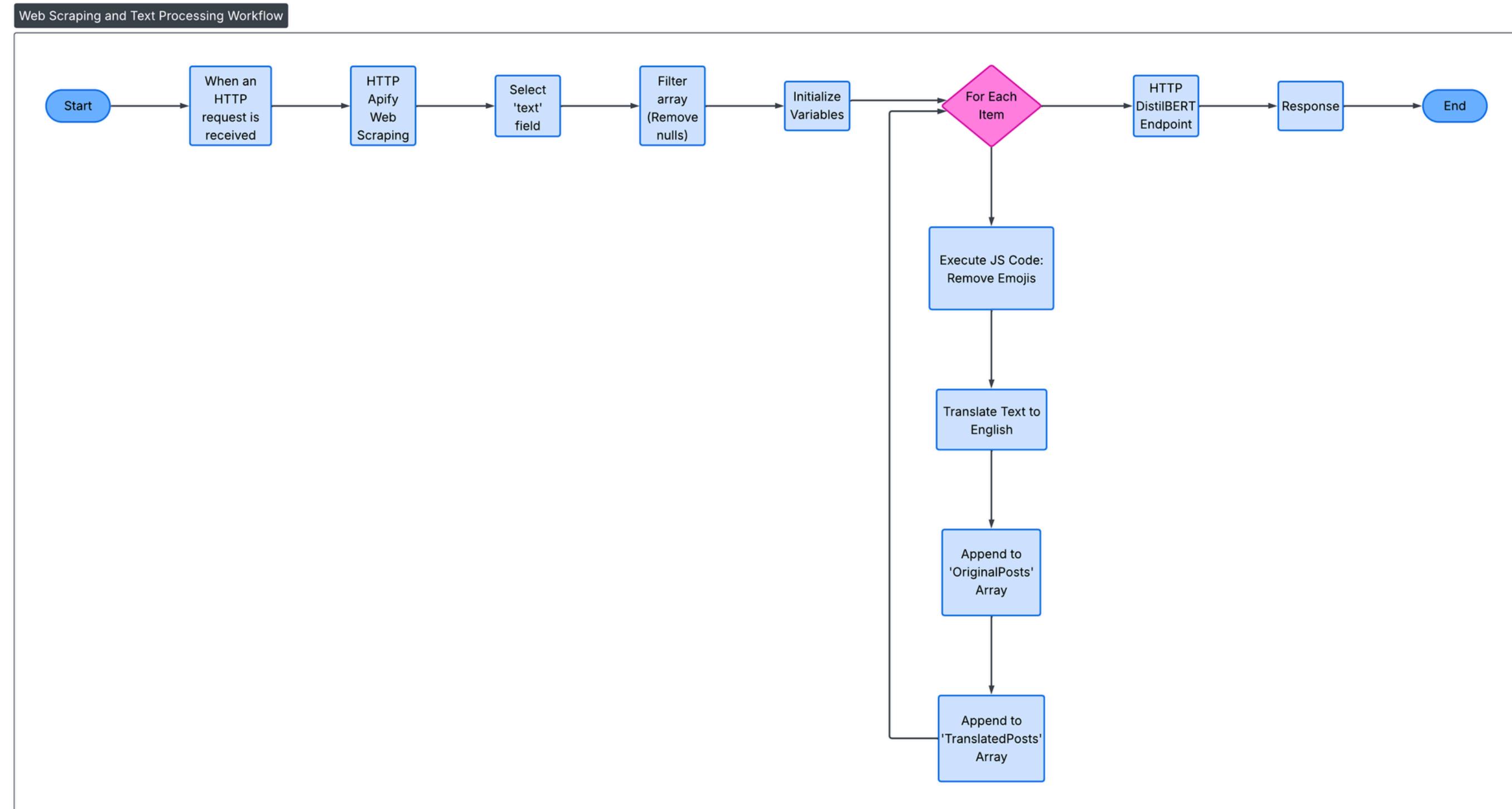
# Deployment



# Deployment

- **Platform:** Deployed on Azure Machine Learning using a Managed Online Endpoint for real-time inference.
- **Workflow:**
  - **Input:** Logic App receives a post URL.
  - **Scoring:** score.py script loads the model and tokenizes text.
  - **Safety:** Inference runs on a secure Standard\_DS2\_v2 instance.
  - **Output:** JSON response with predicted status and probability.

# Deployment



# Conclusion



## Conclusion

- **Bilingual Resource:** Created a custom 2,000-sample dataset in English and Myanmar (Combined\_Data\_2000\_en\_my.csv), addressing the low-resource language gap.
- **Proven Performance:** Successfully implemented a Transformer-based model for effective cross-lingual text classification.
- **Social Impact:** Enables automated monitoring of social-media text to identify potentially harmful mental-health content, supporting SDG 3 (Good Health and Well-Being) through early risk detection and SDG 9 (Industry, Innovation, and Infrastructure) by applying AI-driven NLP systems to social platforms.

# Futurework



## Future Work

- **Data Expansion:** Scale dataset to cover slang, dialects, and both Zawgyi/Unicode fonts.
- **Deployment:** Launch as a real-time Web API or mobile app for instant predictions.
- **Multimodal:** Integrate Speech-to-Text to analyze voice messages (popular in Myanmar).
- **Granularity:** Upgrade from simple sentiment (Positive/Negative) to specific emotion detection (Joy, Anger, Fear).



Thank  
you!

frontier  
tech  
leaders