

EMOTION RECOGNITION

Pipeline for multimedia analysis

PRESENTED BY:

Denys Bespalko, Nikita Orlov and Artjom Musaeljans

Agenda

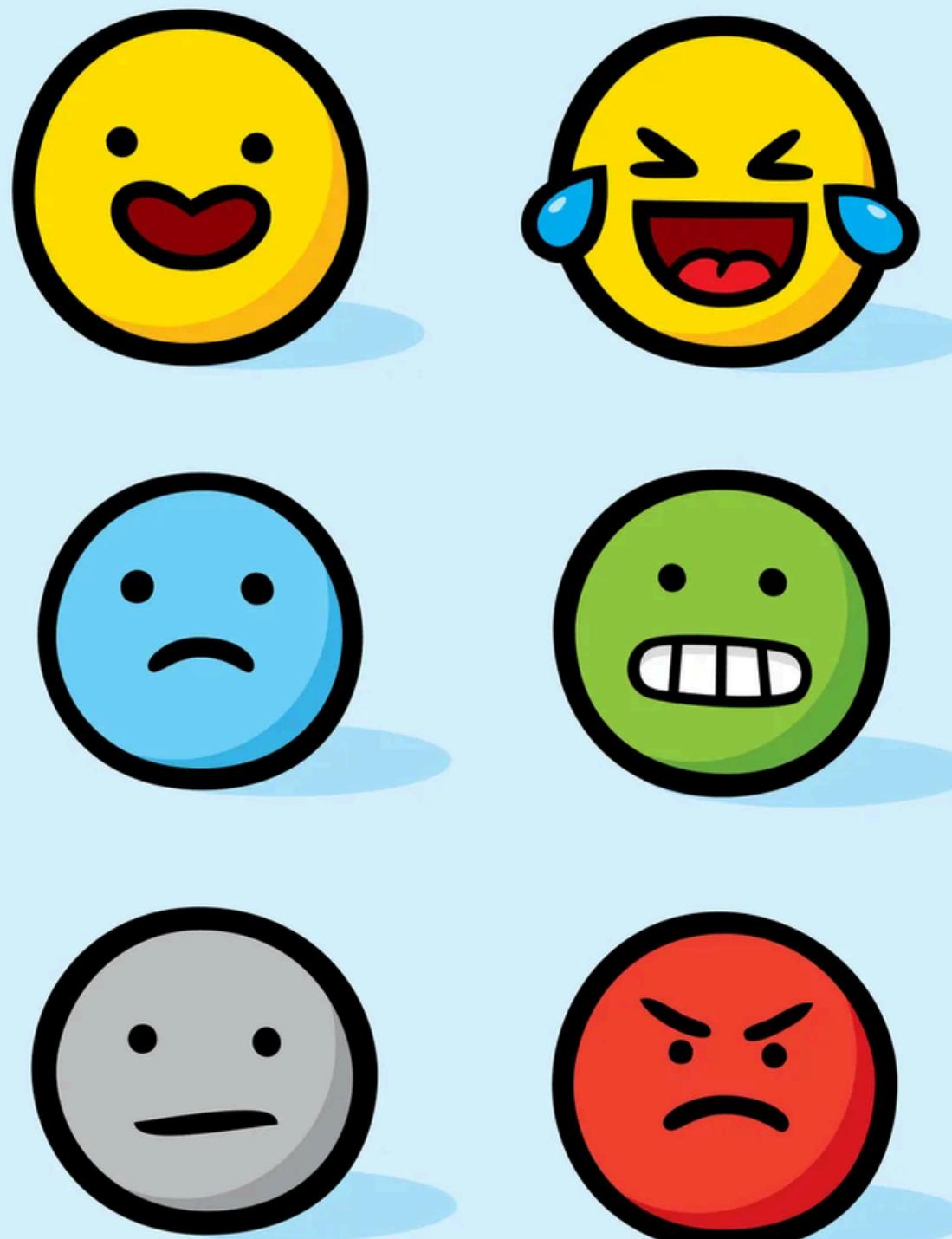
03	Overview
04	Client Value Proposition
05	Data Characteristics & Limitations
06	Model Implementation & Rationale
07	Performance Evaluation & Metrics
08	Ethical Considerations
09	Pipeline Limitations
10	Next Steps & Recommendations

01 Content Intelligence Agency

already has a pipeline that does this, so the idea of our project was to come up with a **language specific** and out-of-the box solution for the problem that could improve it.

02 The pipeline converts audio from videos into transcribed text, translates to English and then classifies to **6 core emotions + neutral** for each segment (Ekman & Friesen, 1971).

03 Project crosses multiple disciplines, such as translation, transcription, sentiment analysis, machine learning and NLP within media context.



Overview

Developed the pipeline designed to analyze multimedia segments using NLP solutions to automate the work of media intelligence analyst.

Client Value Proposition

Emotions Unleashed: Smarter Media Analysis Through AI

Automated Emotion Analysis

Turning video content into meaningful insights.

Enhanced Audience Understanding

Accurately capturing core emotions like happiness, sadness, anger, surprise, fear, and disgust.

Increased Efficiency & Cost Reduction

Reducing manual efforts through automated processing.

Explainable Insights

Transparent AI decisions with advanced Explainable AI techniques.

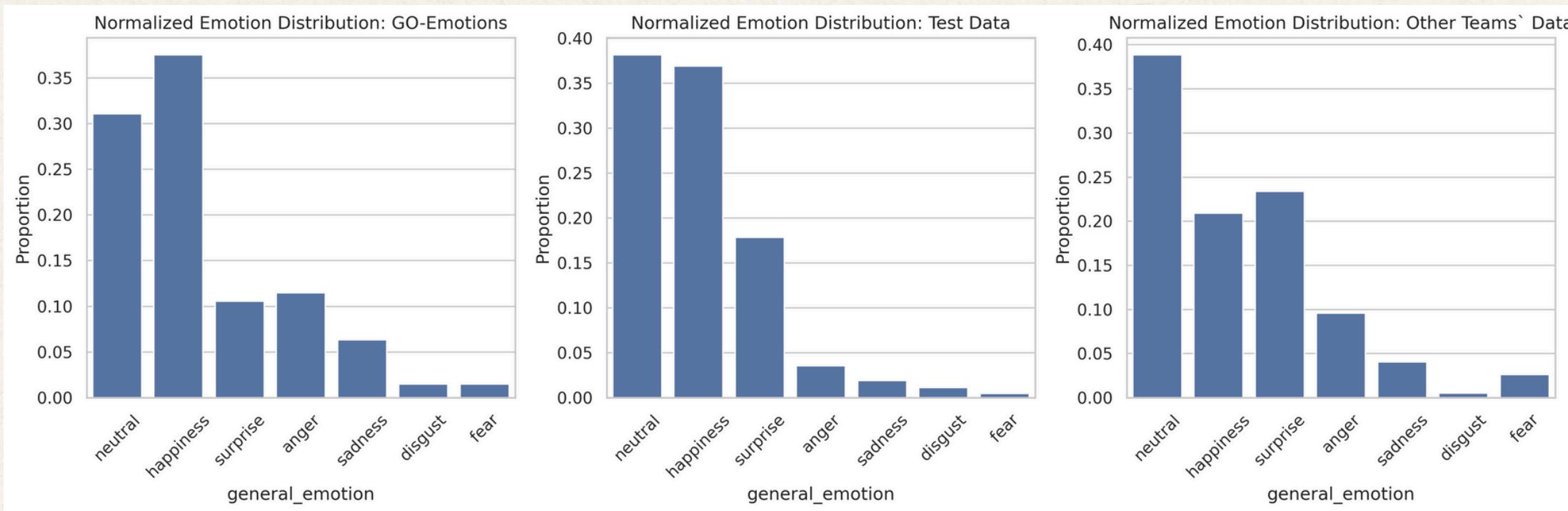
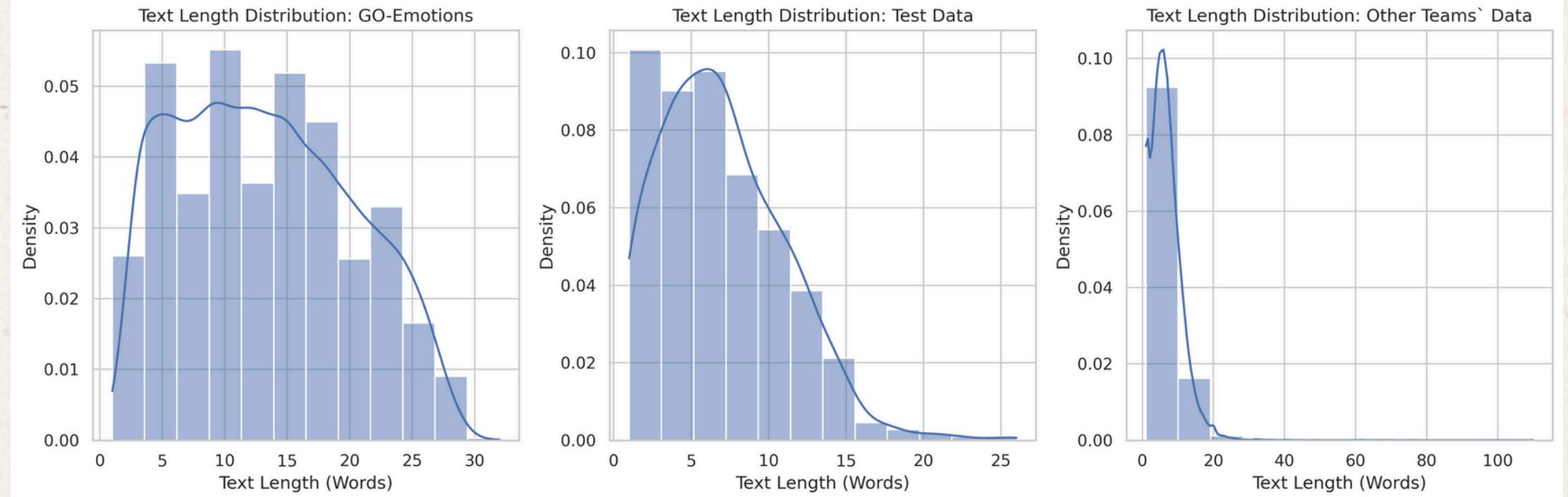
Strategic Advantage

Empowering smarter, data-driven decisions in content creation and marketing.

Data

- **GO-Emotions dataset size: 54257**
- **Other Teams` Data dataset size: 31967**
- **Test Data dataset size: 1043**

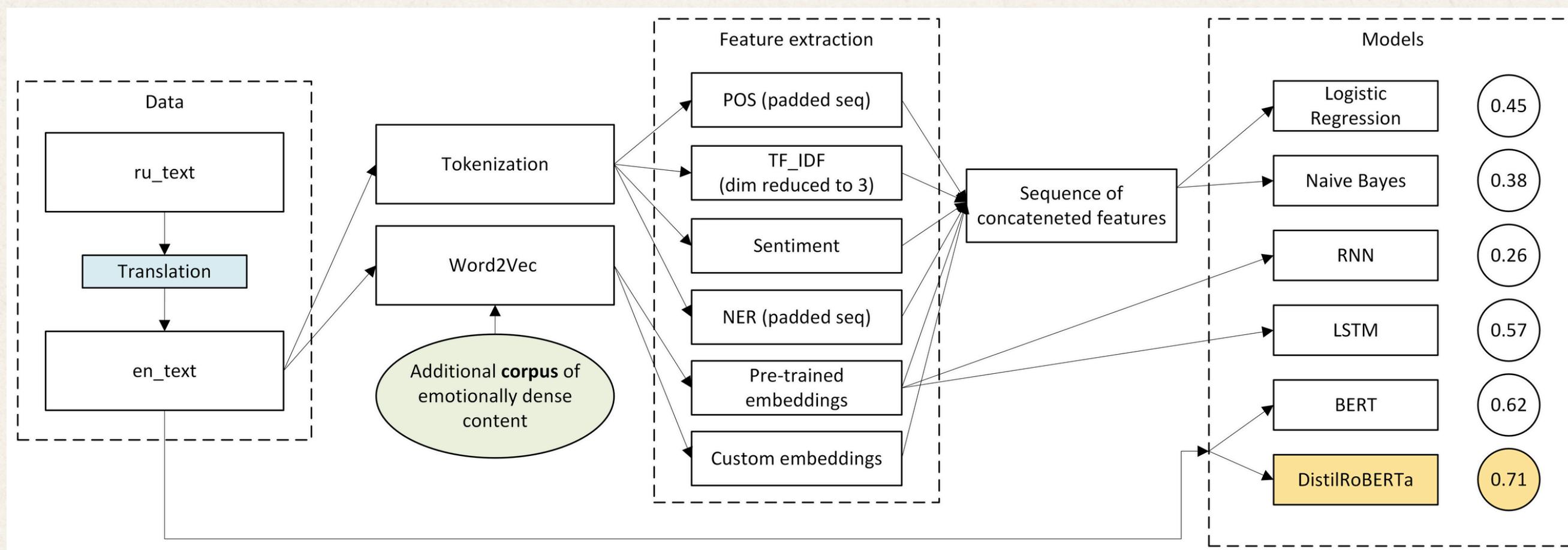
Data was collected from Hugging face or provided by the Content Intelligence Agency current Pipeline



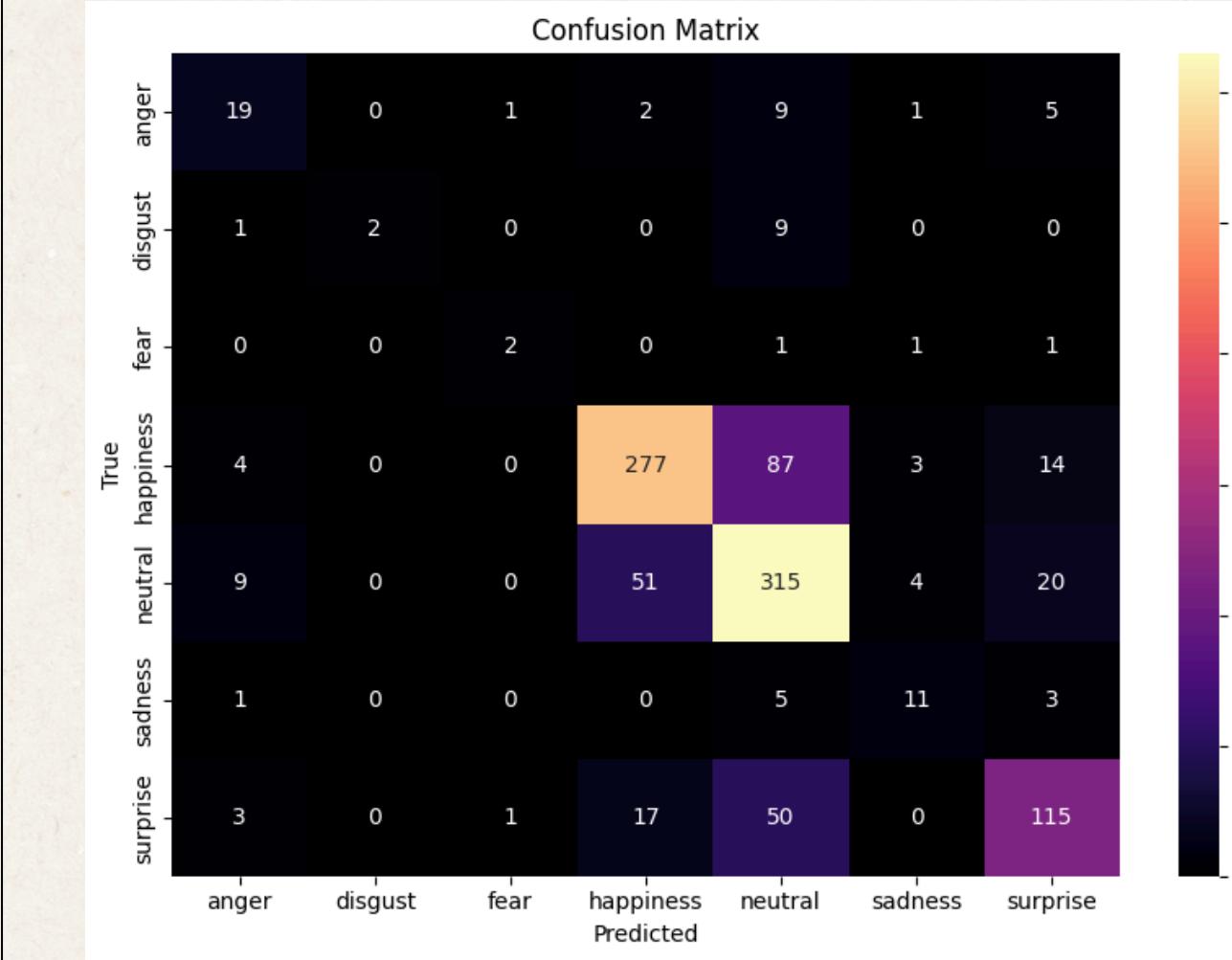
Model

DistillRoBERTa-base was chosen due :

- Light weight and speed
- Pretrained on big English corpus
- Uses attention mechanism to capture context related cues



Performance Evaluation & Metrics



Key Insight
Confusion Matrix reveals consistent misclassifications:

- happiness → neutral (87)
- surprise → neutral (50)
- neutral → happiness (51)

Metric	Value
Accuracy	71%
Macro f1	0.56
Weighted f1	0.71

Strengths

- Strong performance on happiness and neutral
- Good generalization for surprise

Weaknesses

- Low recall for fear, disgust, anger
- Frequent confusion between happiness, neutral, and surprise

Important Trade-off

- Macro F1 > Accuracy
- Balancing performance across all classes, not just frequent ones

Ethical considerations

Key Concerns

- Emotion Misclassification can lead to:
 - Misinterpretation of intent or mood
 - Bias in downstream decision-making (e.g., media analysis, content filtering)
 - Underperformance on Rare Emotions
 - Risk of underrepresenting negative or sensitive emotional states
-

Ethical Design Choices

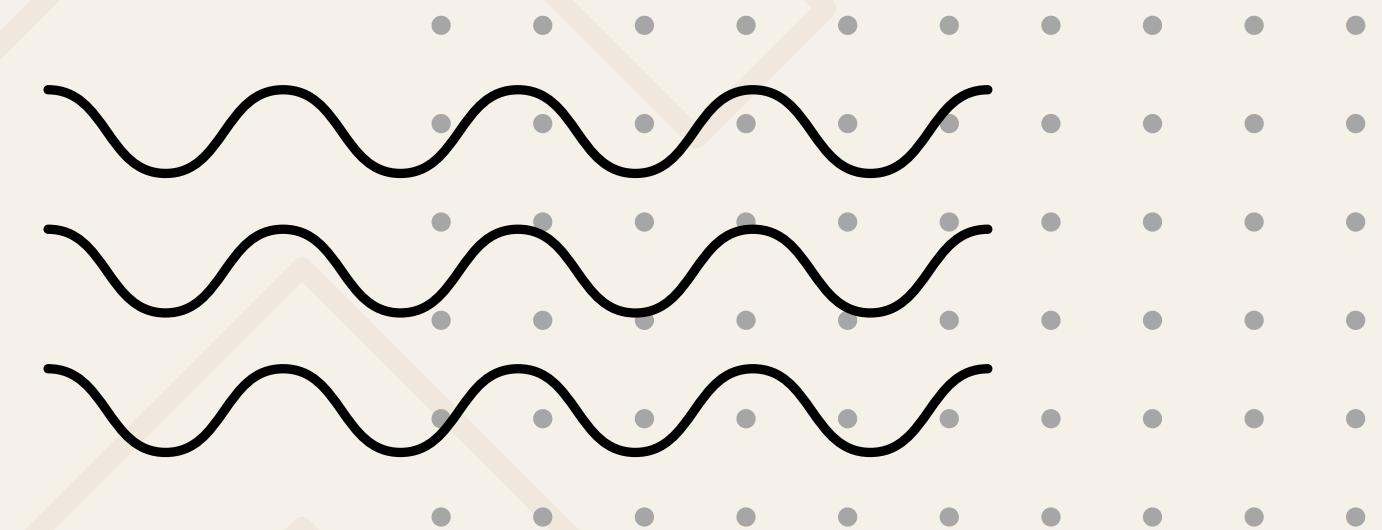
- Prioritized Macro F1 to treat all emotions equally
 - Manual error analysis to expose model biases
 - Used explainable AI (XAI) to verify model focus
-

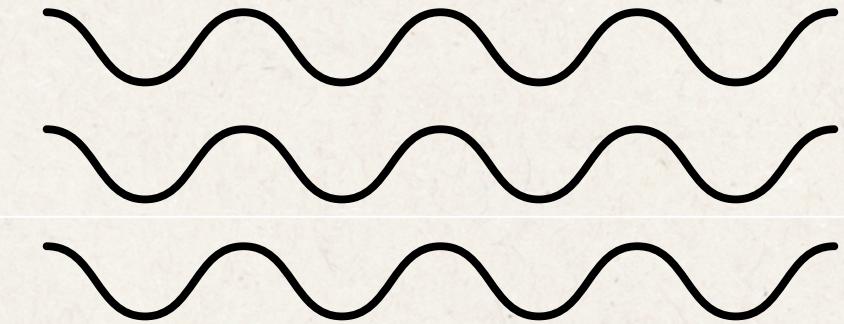
Transparency

- Model card discloses:
 - Data limitations
 - Bias risks
 - Performance disparities across emotions

Pipeline Limitations

Noisy Data	Imbalanced Classes	Limited Context	Interpretability Issues	High Compute Requirements	Ethical Concerns
Inconsistent transcriptions and translations.	Rare emotions (anger, disgust, fear) are underrepresented.	Sentences are processed one by one, losing wider context.	Model often focuses on simple words, not true meaning.	Processing large video data is expensive.	Bias in predictions may lead to misinterpretation.





Next Steps

8

Recommendations

1 Technical Improvements

- Fine-tune on more balanced data for rare emotions
- Explore multi-label classification for mixed emotions
- Integrate audio-based emotion cues (tone, prosody)

2 Recommendations for Client

- Use XAI outputs for validation in high-stakes cases
- Deploy model for offline analysis, not real-time decisions
- Include a “low confidence” flag for manual review

3 Continuous Iteration

- Update pipeline with new labeled data
- Regularly audit class-wise performance