NLP INSIGHTS FROM INSTAGRAM POSTS ON MPOX

Presented by;

1. Judah Anangwe
2. Doreen Kahare
3. Yvonne Karinge

Date: 16/07/2025

INTRODUCTION

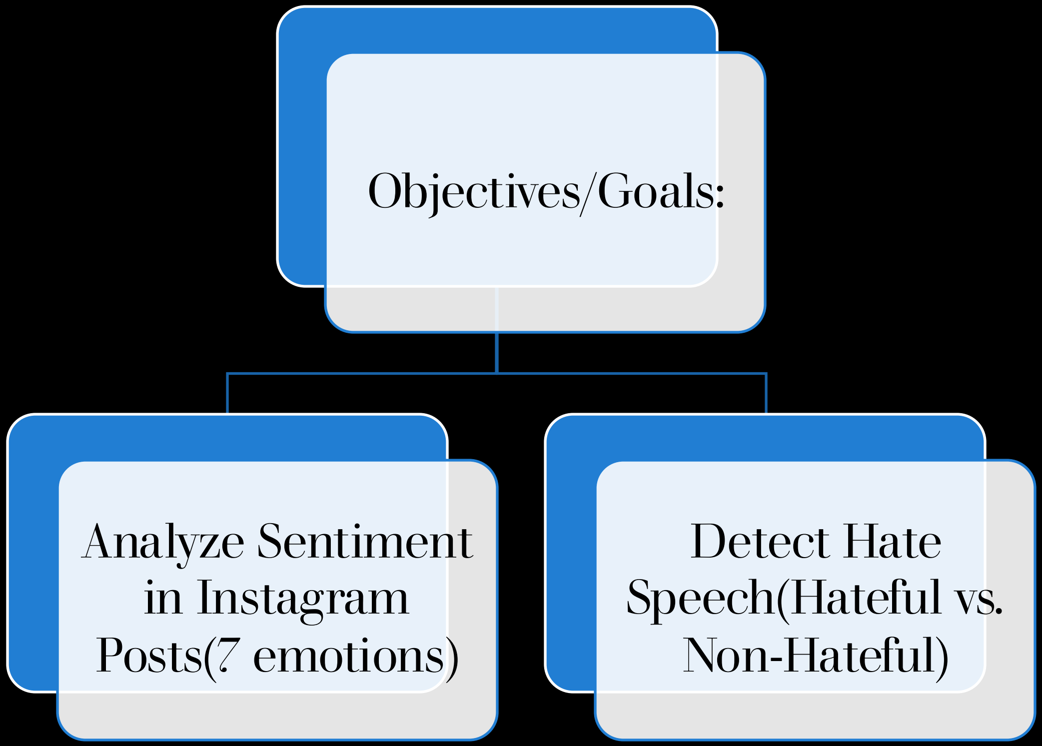
In this project, we aim to leverage NLP techniques to analyze user-generated Instagram content related to Mpox(formerly known as Monkeypox)

Why Analyze Instagram Posts on MPOX?

- Instagram became a hub of emotions during the MPOX outbreak
- Posts included fear, questions and misinformation
- Understanding this helps improve health messaging

PROJECT OVERVIEW

- *Analyze Instagram Posts using Natural Language Processing(NLP)*
- *Understand Public Sentiment & Detect Online Toxicity*



Data Understanding

Dataset: Mpox Instagram Dataset – Sentiment and Hate Analysis was sourced from Kaggle.

https://www.kaggle.com/datasets/thakurnirmalya/mpox-instagram-dataset-sentiment-and-hate-analysis

There were 59299 Instagram captions

Key Features: Translated Post Description, Sentiment, Hate & Language

No Missing Values

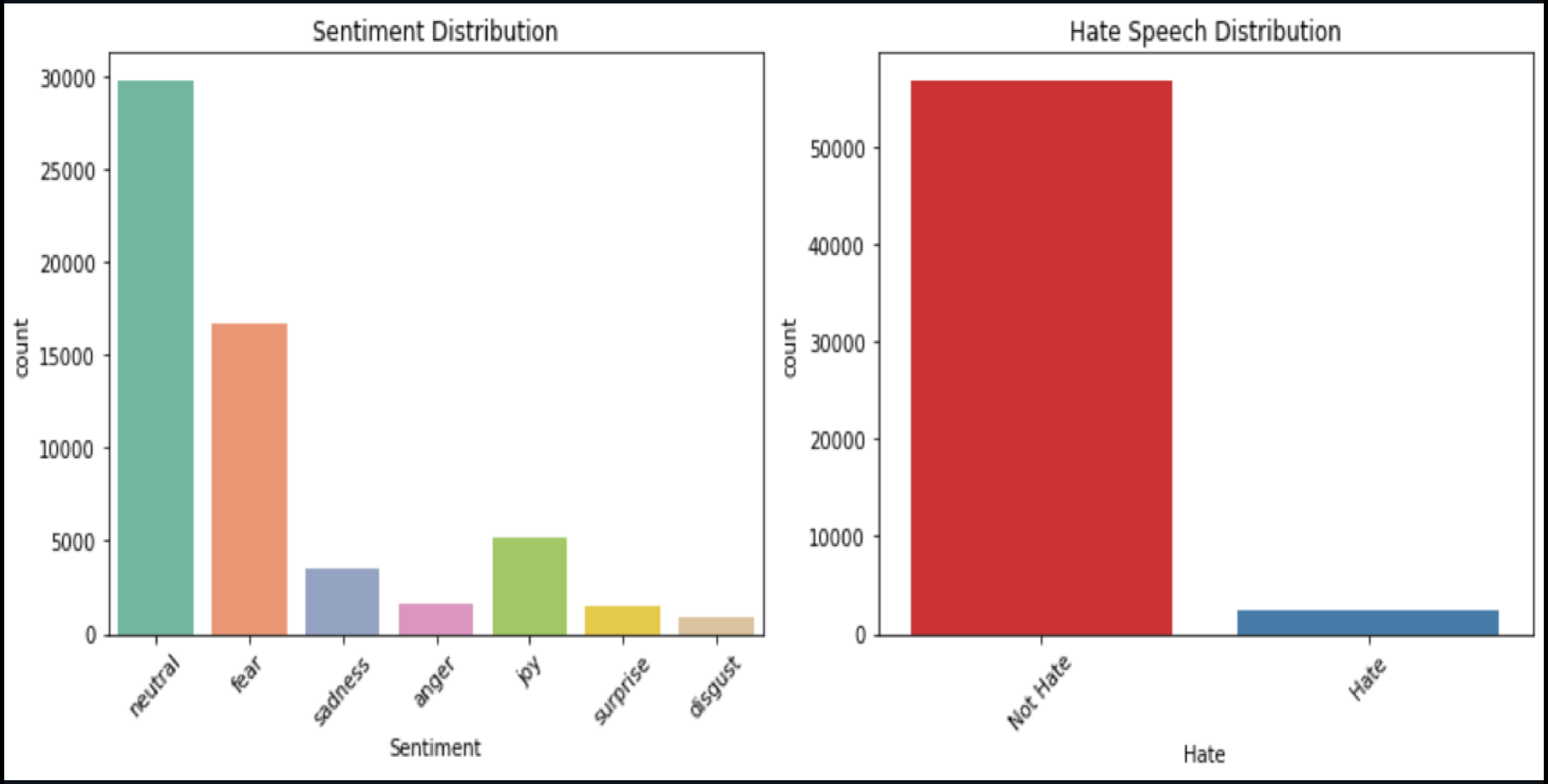
97% of post in English

4.25% flagged as Hate Speech

- Initially we had decided to use the top 3 languages
- However, we decided to exclude the non-English posts during data cleaning and preparation

Language Code	Language	Number of Posts
`en`	English	58,616
`fi`	Finnish	566
`es`	Spanish	128
`id`	Indonesian	113
`fr`	French	98
`it`	Italian	69
`af`	Afrikaans	62
`pt`	Portuguese	47
`no`	Norwegian	46
`sv`	Swedish	37

Sentiment & Hate Speech Distribution



Takeaway:

- ❖ Dominant emotions: fear and neutral
- ❖ Hateful posts often linked to fear/misinformation

NLP TECHNIQUES USED

1. Tokenization and Lemmatization:

Definition:

- **Tokenization** is the process of dividing a text into individual units (tokens). Its purpose is to convert raw text into a format that computers can understand and process, forming the basis for further analysis.
- **Lemmatization** reduces words to their base or dictionary form (lemma), taking into account the word's part of speech and context. Its purpose is to normalize words, allowing for better analysis and understanding of the text.

Removed special characters & numbers like emojis, links etc

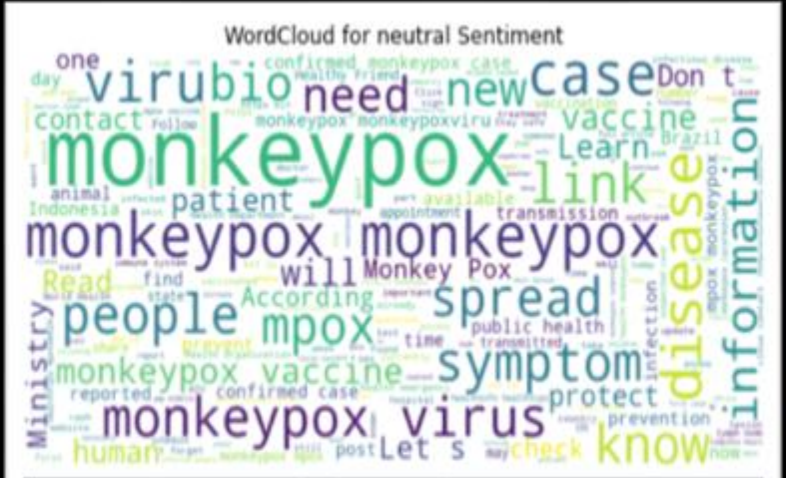
Text standardized (lowercase, no punctuation)

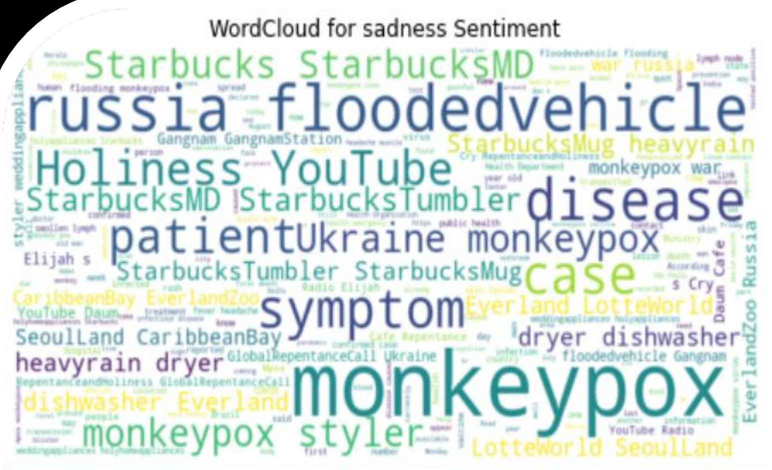
Tokenized the text

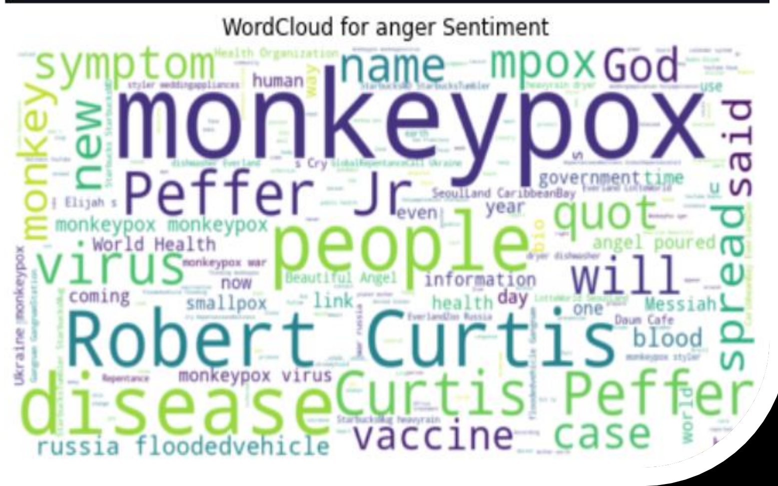
Removed stopwords

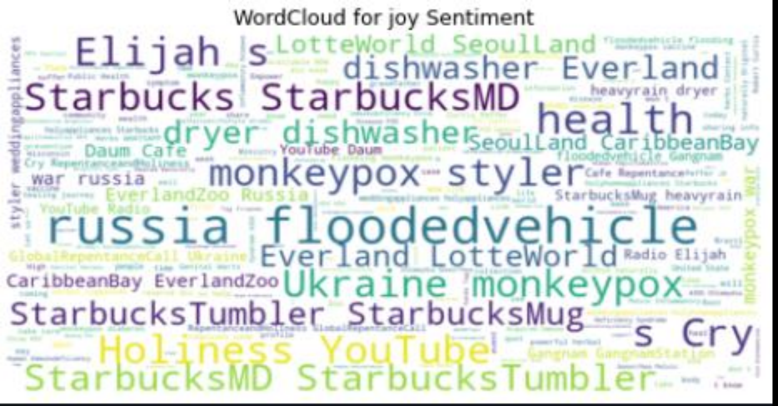
Applied preprocessing and obtained text_cleaned data

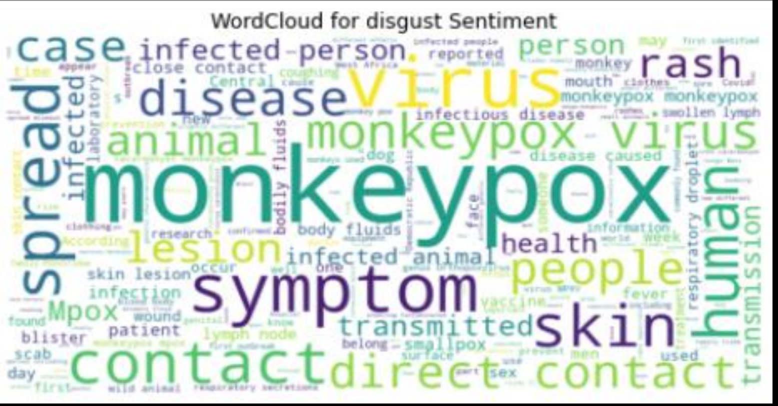
2. Word Clouds & Frequency Analysis





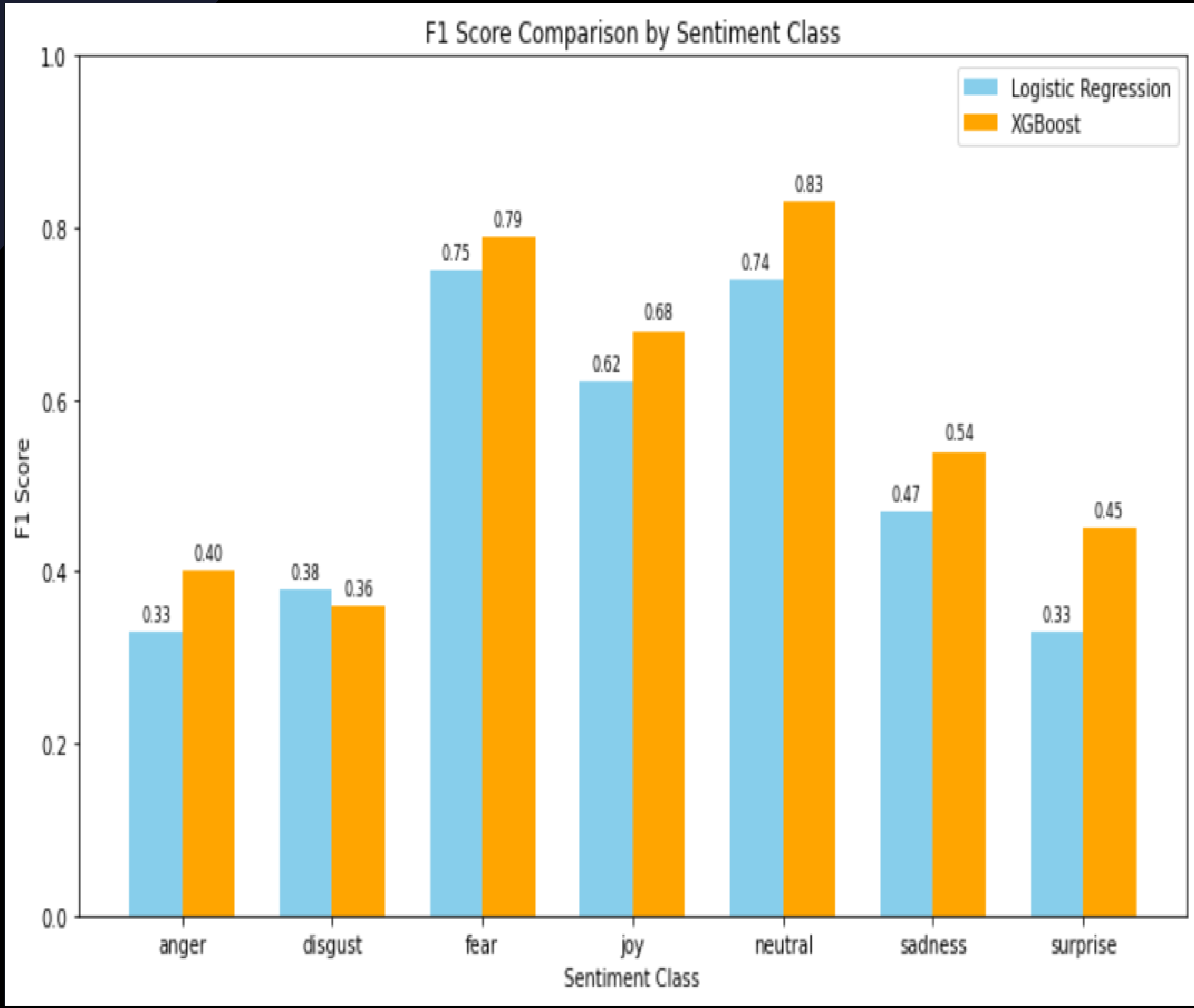






❖ These visuals gives us a summary of the most frequent words in each sentiment category

3. Deep Learning for Sentiments



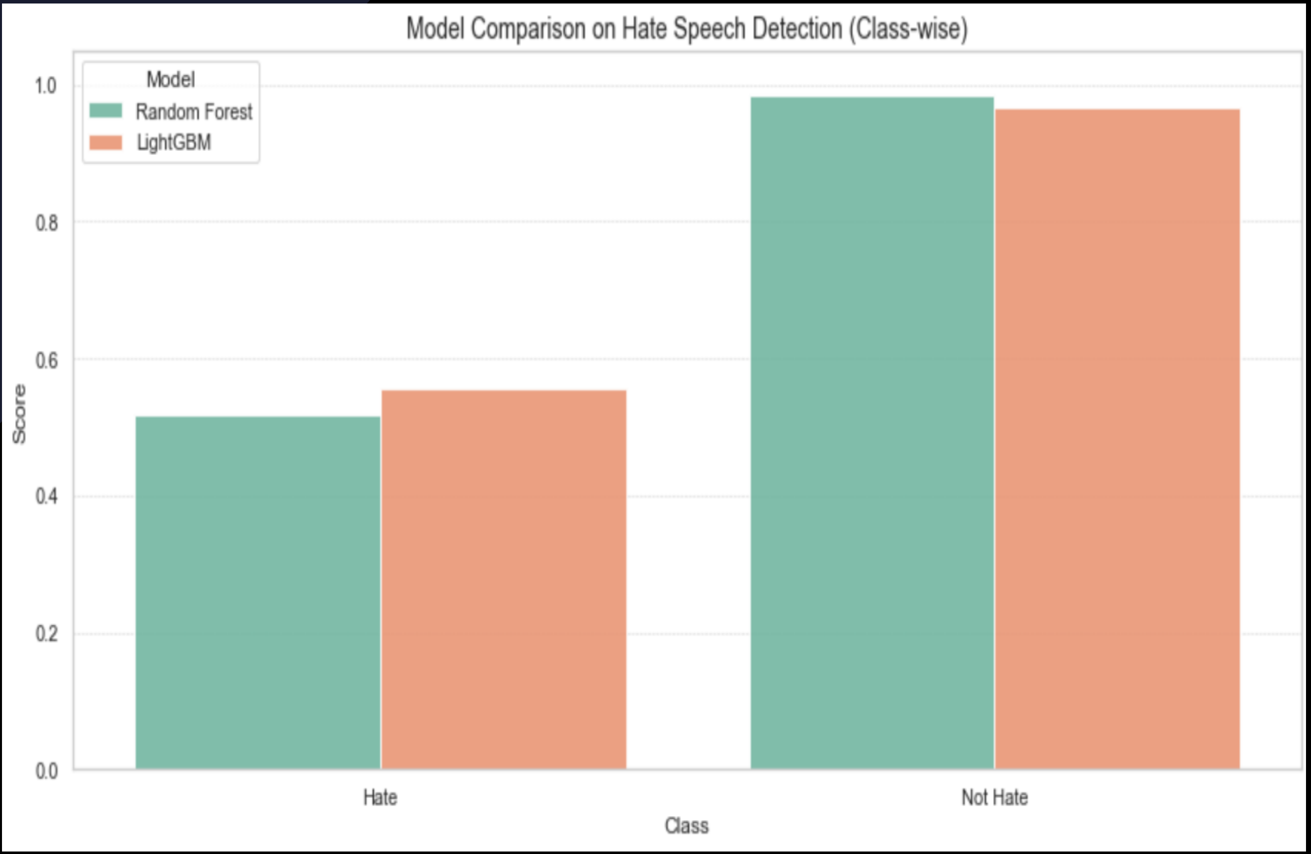
Sentiment Analysis Models Comparison

- Model 1: **Logistic Regression**
  - Tuned parameters: C=10, max_features=5000, ngram_range=(1,1)
  - Accuracy: 67%
- Model 2: **XGBoost Classifier**
  - Tuned parameters: learning_rate=0.3, max_depth=6, n_estimators=200, max_features=5000, ngram_range=(1,1)
  - Accuracy: 77%

Metric	Logistic Regression	XGBoost
Accuracy	0.67	0.77
Macro Avg F1	0.52	0.58
Weighted Avg F1	0.69	0.76

Performance insights:
Neutral & Fear were well-predicted by both models
XGBoost achieves excellent precision & recall f(0.76/0.91) and fear (0.81/0.77)
Joy and Sadness perform moderately
Anger, Disgust and Surprise show low recall across both models due to class imbalance/ less distinctive language patterns

4. Binary Classification for Hate Speech



MODEL 1:

- ✓ Random Forest is biased toward the majority class(Not Hate), leading to poor recall for hate speech.
- ✓ Severely underperforms on the minority class(Hate), despite high overall accuracy

MODEL 2:

- ➢ LightGBM provides much better performance on the minority class(Hate), making it more suitable for this task
- ➢ It handles class imbalance better and improves hate detection without sacrificing much accuracy

Best/Preferred model: **LightGBM**
We'll use LightGBM for hate speech classification especially since detecting hate speech(minority class) is our primary goal

Conclusion

PEOPLE'S RESPONSE: ANALYZING HOW INDIVIDUALS EXPRESS THEIR SENTIMENTS IN SOCIAL MEDIA POST

SENTIMENT TRENDS: SUPPORT EFFECTIVE CRISIS COMMUNICATION

PUBLIC ENGAGEMENT: COMBINED SENTIMENT & HATE SPEECH = SMARTER OUTREACH

Challenges & Limitations

1. **Class Imbalance**: Imbalanced class distribution led to lower recall in minority emotion categories
2. **Language Scope**: Only English data was modeled due to data sparsity in other languages
3. **Topic Sensitivity**: The overlap of public health discourse with potentially sensitive terms required careful handling to avoid false positives in hate speech detection

Recommendation

Deploy	Deploy models for real-time monitoring of social media content
Implemen t	Implement continuous learning to adapt to evolving language patterns
Use	Use model explanations for content moderation decisions
Monitor	Monitor model performance across different languages and time periods

Future Work

01

Experiment with transformer-based models(BERT, RoBERTa)

02

Implement ensemble methods for improved robustness

03

Add multilingual support for non-English content

04

Develop real-time prediction API

THANK YOU!

- Questions?