Sentiment Analysis of Tweets on Apple and Google Products

September 2025

Presented by Group 5

Technical Mentor: George Kamundia

Introduction

Project Summary & Business Understanding

Business Problem

Apple and Google constantly face public scrutiny on social media. Understanding real-time customer sentiment is crucial to improve products, marketing, and customer satisfaction.

Core Question

"Can we automatically classify the sentiment of tweets about Apple and Google products to support actionable business insights?"

Project Objectives

- ❑ Determine overall public sentiment towards Apple and Google products
- ❑ Identify sentiment drivers in tweets
- ❑ Provide actionable insights for business decisions

Datas et Overview

The dataset contains over **9,093 tweets** about Apple and Google products, each labeled with a sentiment.

9,093

Total Records

3

Features

22

Duplicates

Data Columns

- **tweet_text**: The content of the tweet (0.01% missing)
- **emotion_in_tweet_is_directed_at**: Brand/product target (63.8% missing)
- **is_there_an_emotion_directed_at_a_brand_or_product**: Sentiment label

Data Cleaning

- **Column Standardization**

Shortening names for easy reference

- **Removing duplicates**

Preventing duplicate samples from overweighting certain classes hence every unique tweet target pair appears only once.

- **Mapping sentiment unique values**

Consolidates the neutral emotions and remove emotions from the other parameter

Brand Mentions Analysis

66%

More Mentions

Apple has 66% more emotion mentions (3,834) than Google (2,309)

5x

Brand vs Product

Corporate brands mentioned about 5 times more often than specific products

85%

Ecosystem Gap

Apple ecosystem has 85% more emotion mentions than Google ecosystem

Strategic Implications

- ❑ Apple generates stronger emotional engagement across its ecosystem
- ❑ Corporate branding drives more emotional discourse than individual products
- ❑ App experience is a major emotional driver, particularly for Apple (5.8x more mentions)
- ❑ Opportunity for Google to enhance emotional engagement with its wider product range

Word Cloud Visualization by Sentiment

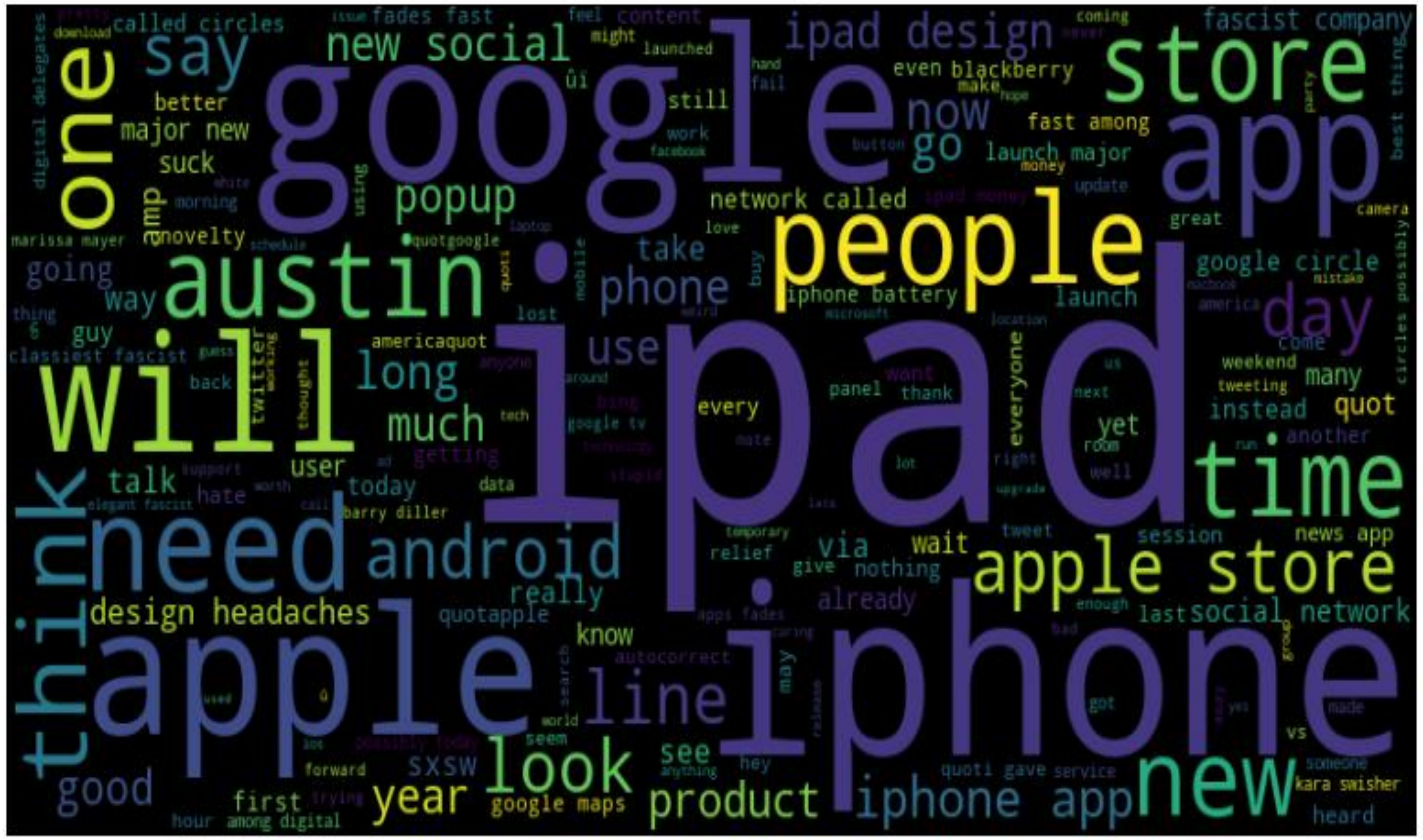
Positive Sentiment

Most frequent words: "apple", "google", "store", "app", "ipad", "sxsw", "new"

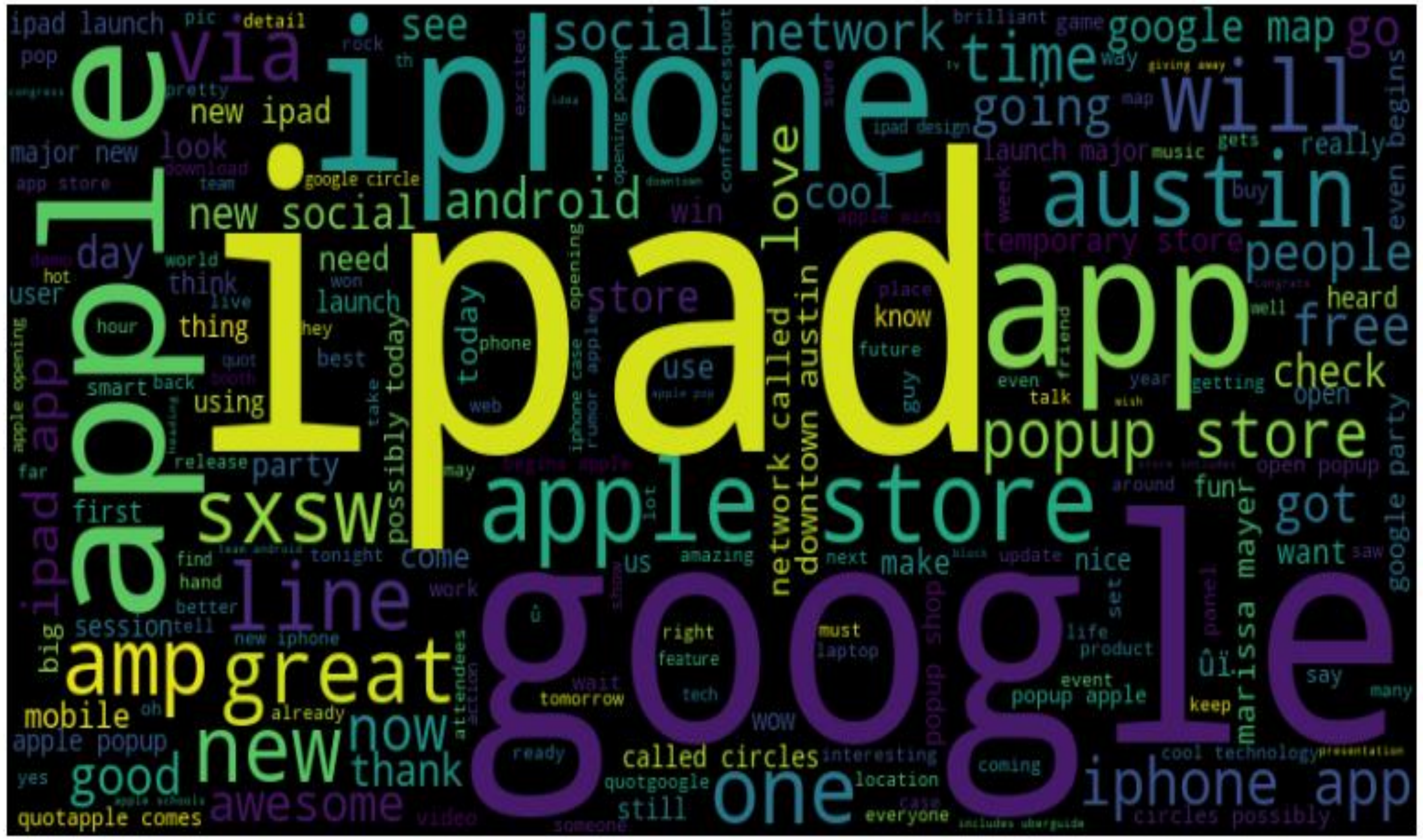
Negative Sentiment

Most frequent words: "apple", "google", "iphone", "ipad", "app", "store"

Word Cloud for Negative



Word Cloud for Positive



Neutral Sentiment

Most frequent words: "apple", "google", "ipad", "sxsw", "iphone", "store", "app", "new", "rt"

This visualization helps identify key terms driving different emotional responses.

Key Findings

Overall Distribution Characteristics

❑ Central tendency and variability are highly consistent across sentiment classes

Central Tendency & Variability Patterns

❑ Mean word counts show minimal variation across sentiments (range: 14.1 to 15.9
❑ words)
❑ Medians align closely with means, confirming symmetrical distributions with minimal skew

Outlier Detection Results

❑ Low-Severity Outliers: Limited outliers present, primarily in Neutral class

Distribution Concentration Patterns

❑ Data density follows similar patterns across all sentiment categories
❑ Violin plots show consistent distribution shapes with minimal class differentiation

Behavioral Pattern Recognition

❑ No significant correlation found between sentiment polarity and message length in
❑ this dataset
❑ Customers maintain similar text lengths regardless of emotional context

Statistical Validation

❑ Quantitative analysis confirms visual observations with precise numerical metrics

Data Preprocessing

❑ **Tokenization**

Split text into word/tokens

❑ **Stop word removal**

Removing words that will add little sentiment information.

❑ **Lemmatization**

Reducing words to their original form for consistency

❑ **Vectorization**

Converting text to numerical features.

Our Modeling Strategy

Feature Engineering

TF-IDF Vectorization converts tweets into numerical data, capturing word importance relative to the corpus. Maximum 5,000 features to focus on relevant terms.

Handling Class Imbalance

SMOTE generates synthetic samples for the minority 'Negative' class, creating a balanced dataset and preventing model bias.

Model Training & Evaluation

Testing various algorithms from simple baselines to complex models. Performance measured using Accuracy, Precision, Recall, and F1-Score.

Final Model Evaluation

Model	Train Accuracy	Test Accuracy	Test F1 (Weighted)	Test F1 (Macro)	Precision (Weighted)	Recall (Weighted)	Training Time (s)	Prediction Time (s)	Overfitting Score
Random Forest (Tuned)	0.97	0.677	0.673	0.584	0.671	0.677	291.4	3.909	0.292
XGBoost	0.805	0.66	0.652	0.526	0.648	0.66	39.35	0.298	0.145
Neural Network	0.954	0.649	0.648	0.553	0.647	0.649	224.73	0.051	0.306
Logistic Regression (Tuned)	0.956	0.632	0.637	0.548	0.643	0.632	1765.67	0.005	0.324
Naive Bayes (Tuned)	0.826	0.596	0.613	0.529	0.65	0.596	0.75	0.004	0.23
Naive Bayes (Untuned)	0.803	0.574	0.595	0.511	0.653	0.574	0.01	0.006	0.23

Selected Model: Tuned XGBoost Classifier

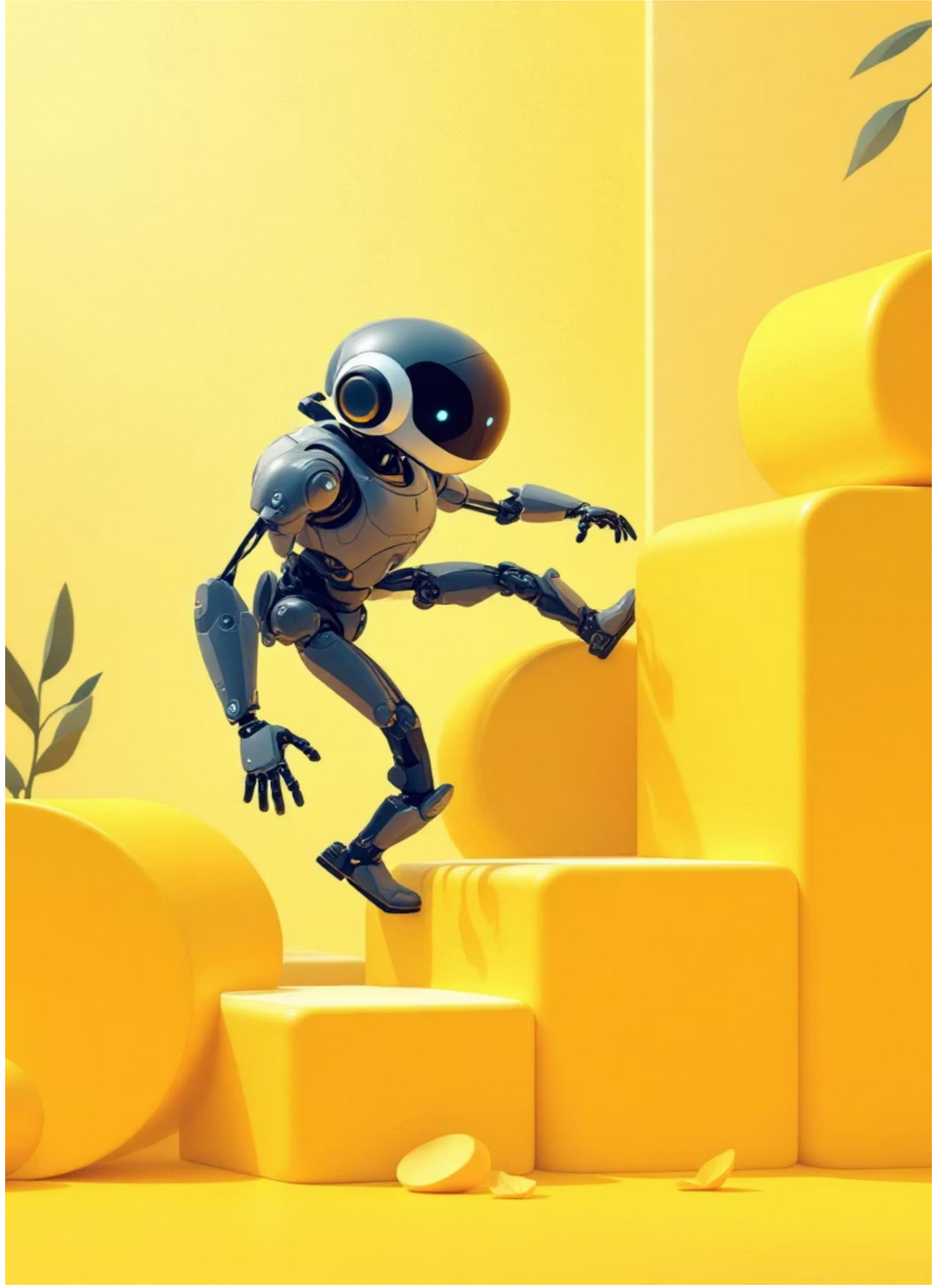
- ❑ Demonstrated the best balance of performance
- ❑ Generalization with the lowest overfitting score (14.5%)
- ❑ Competitive test accuracy (66.9%).

Business Recommendation

- ❑ **Deploy XGBoost Model**
- ❑ **Monitor Sentiment Trends Over Time**
- ❑ **Implement Periodic Model Retraining**
- ❑ **Expand Analysis and Exploration**

Limitations

- ❑ Class Imbalance
- ❑ Short Text Nature of Tweets
- ❑ Dynamic Language and Slang
- ❑ Model Generalization on Current Events
- ❑ Performance Ceiling of Traditional ML



Thank you

Q & A?