Sentiment Analysis of Tweets on Apple and Google Products

September 2025

Presented by Group 5

TM: George Kamundia

Business Understanding

- Social media plays a huge role in shaping brand perception
- Apple and Google face constant public scrutiny about their products.

Business problem

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

Stakeholders

- ***Apple & Google Executives*** – for product strategy.
- ***Marketing Teams*** – to refine campaigns.
- ***Product Managers*** – to address customer concerns.
- ***Customer Support*** – to monitor satisfaction trends.

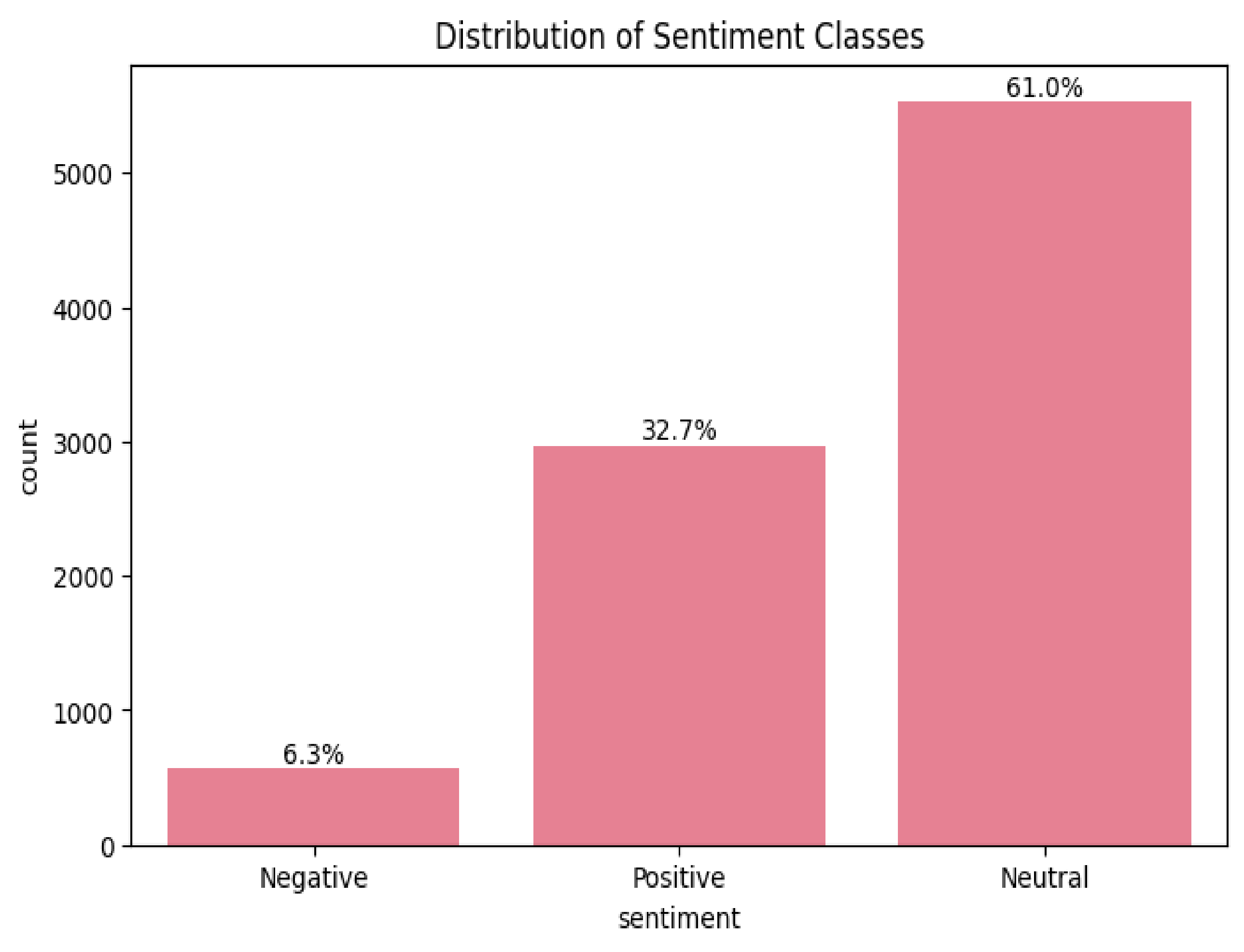
Objectives

- Understand overall public sentiment (positive, neutral, negative).
- Identify drivers of sentiment for Apple vs. Google.
- Provide actionable insights to improve marketing & product development.

Data Understanding

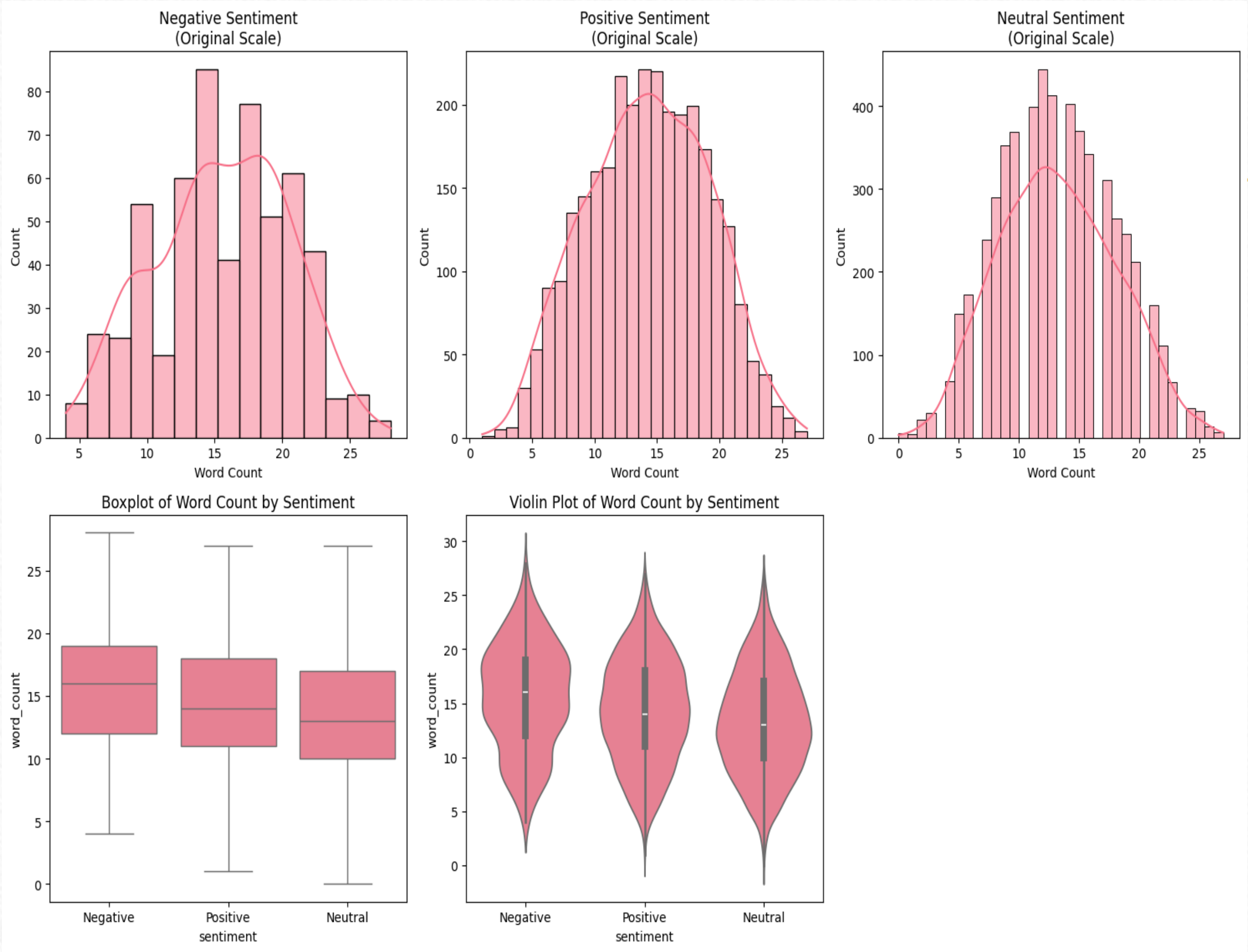
- **Source**: 9,000+ tweets labeled by sentiment.
- **Sentiment distribution**:
  - Neutral (61%)
  - Positive (33%)
  - Negative (6%)
- Apple tweets generated 85% more mentions than Google.

Visualizations (EDA)



- **Neutral:** 61% of tweets
- **Positive:** 32.7% of tweets
- **Negative:** 6.3% of tweets

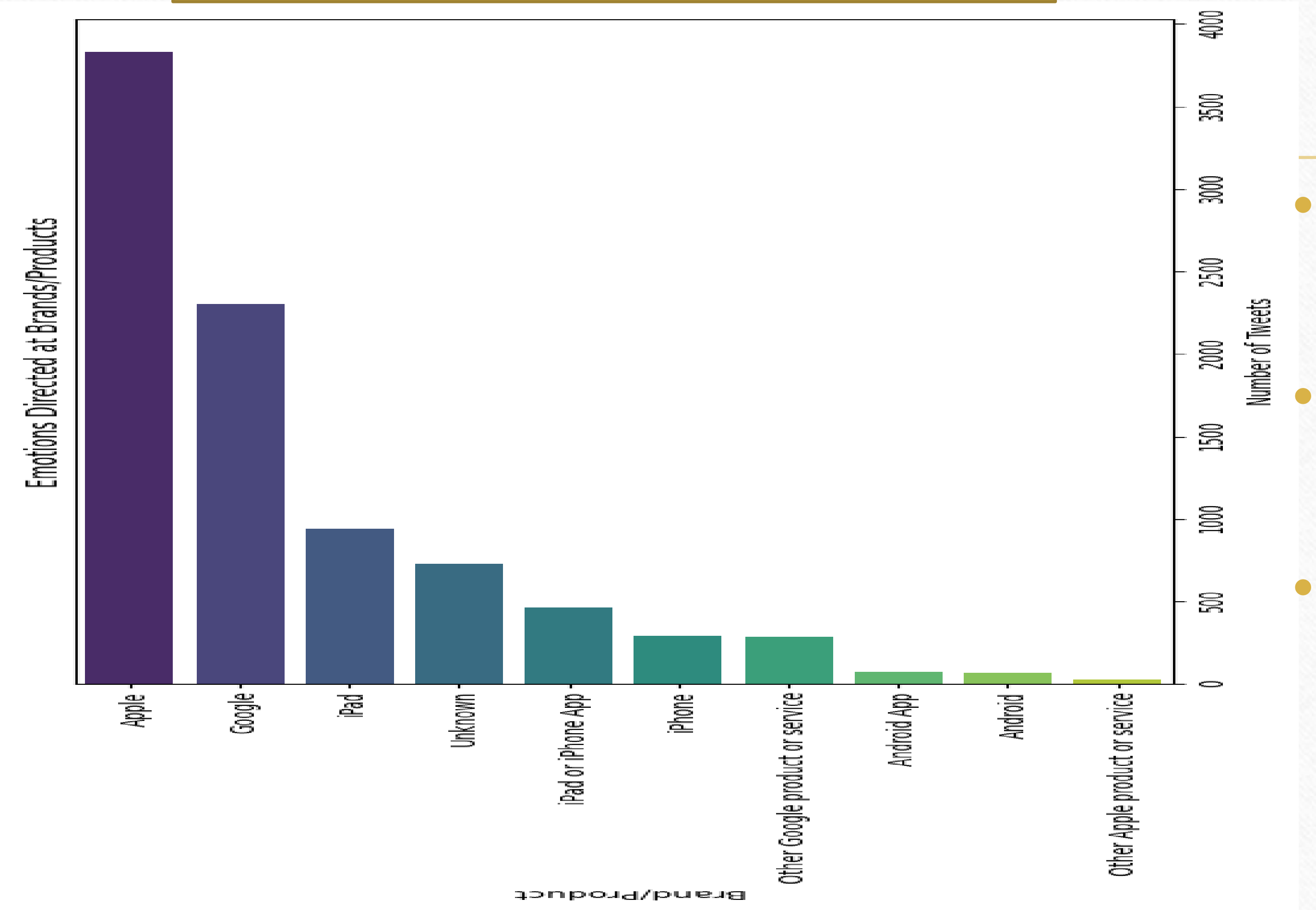
Visualizations (EDA) Cont.....



- Overall Distribution Characteristics
- Central Tendency & Variability Patterns
- Outlier Detection
- Distribution Concentration Patterns
- Behavioral Pattern Recognition
- Statistical Validation

Visualizations (EDA) Cont…..

Emotions Directed at Brands



- Apple generates stronger emotional engagement across its ecosystem
- Google has opportunity to increase emotional engagement with its broader product suite
- Corporate branding matters more than individual product names in emotional discourse

Data Preparation

- Cleaning tweets: removed duplicates, URLs, mentions, emoji's, and irrelevant text.
- Standardized sentiment into three categories: Positive, Negative, Neutral.
- Imputed missing brand mentions (Apple/Google).
- Mapped products into ecosystems: Apple vs Google.

Modeling

- Tested multiple models
- Addressed class imbalance with **SMOTE**.
- Measured performance using Accuracy, Precision, Recall, F1-Score.

Baseline Models

- Naïve Bayes
- Logistic Regression.

Advanced Models

- Random Forest
- XGBoost
- Neural Networks

Evaluation

Model	Train_Accuracy	Test_Accuracy	Test_F1_Weighted	Test_F1_Macro	Test_Precision_Weighted	Test_Recall_Weighted	Training_Time(s)	Prediction_Time(s)	Overfitting_Score
XGBoost	0.805	0.66	0.652	0.526	0.648	0.66	524.43	3.556	0.145
Naive Bayes (Untuned)	0.803	0.574	0.595	0.511	0.653	0.574	0.01	0.006	0.23
Naive Bayes (Tuned)	0.826	0.596	0.613	0.529	0.65	0.596	53.02	0.004	0.23
Random Forest (Tuned)	0.97	0.677	0.673	0.584	0.671	0.677	902.66	7.69	0.292
Neural Network	0.954	0.649	0.648	0.553	0.647	0.649	743.36	0.088	0.306
Logistic Regression (Tuned)	0.956	0.632	0.637	0.548	0.643	0.632	4098.08	0.014	0.324

Evaluation Cont….

- **XGBoost** best model because it offered the most reliable balance between performance, efficiency, and generalization.
- **Random Forest (Tuned)** achieved the highest test accuracy (0.677) but showed higher overfitting compared to XGBoost.
- **Neural Network** and **Logistic Regression (Tuned)** achieved good training scores but struggled with generalization, showing signs of overfitting.
- **Naive Bayes (Untuned/Tuned)** was the fastest to train and predict, but its accuracy and F1 scores were noticeably lower.
- Best trade-off between performance and generalization → XGBoost
- Fastest → Naive Bayes
- Highest raw accuracy → Random Forest (but more overfitting)

Model Selection

Model	Train_Accuracy	Test_Accuracy	Test_F1_Weighted	Test_F1_Macro	Test_Precision_Weighted	Test_Recall_Weighted	Training_Time(s)	Prediction_Time(s)	Overfitting_Score
XGBoost	0.882	0.675	0.666	0.546	0.662	0.675	1096.29	0.324	0.207
Random Forest	0.97	0.677	0.673	0.584	0.671	0.677	150.11	2.305	0.292
Neural Network	0.939	0.619	0.626	0.542	0.64	0.619	832.11	0.061	0.321

XGBoost offers more balanced performance and robustness, therefore selected for deployment.

Tuning of the top three models for selection:

- XGBoost
- Random Forest
- Neural Network
- **XGBoost** demonstrated balanced performance with relatively low overfitting
- **Random Forest** achieved the highest test accuracy and F1-score, but higher overfitting score compared to XGBoost
- **Neural Network u**nderperformed compared to tree-based models, with lower accuracy and highest overfitting.

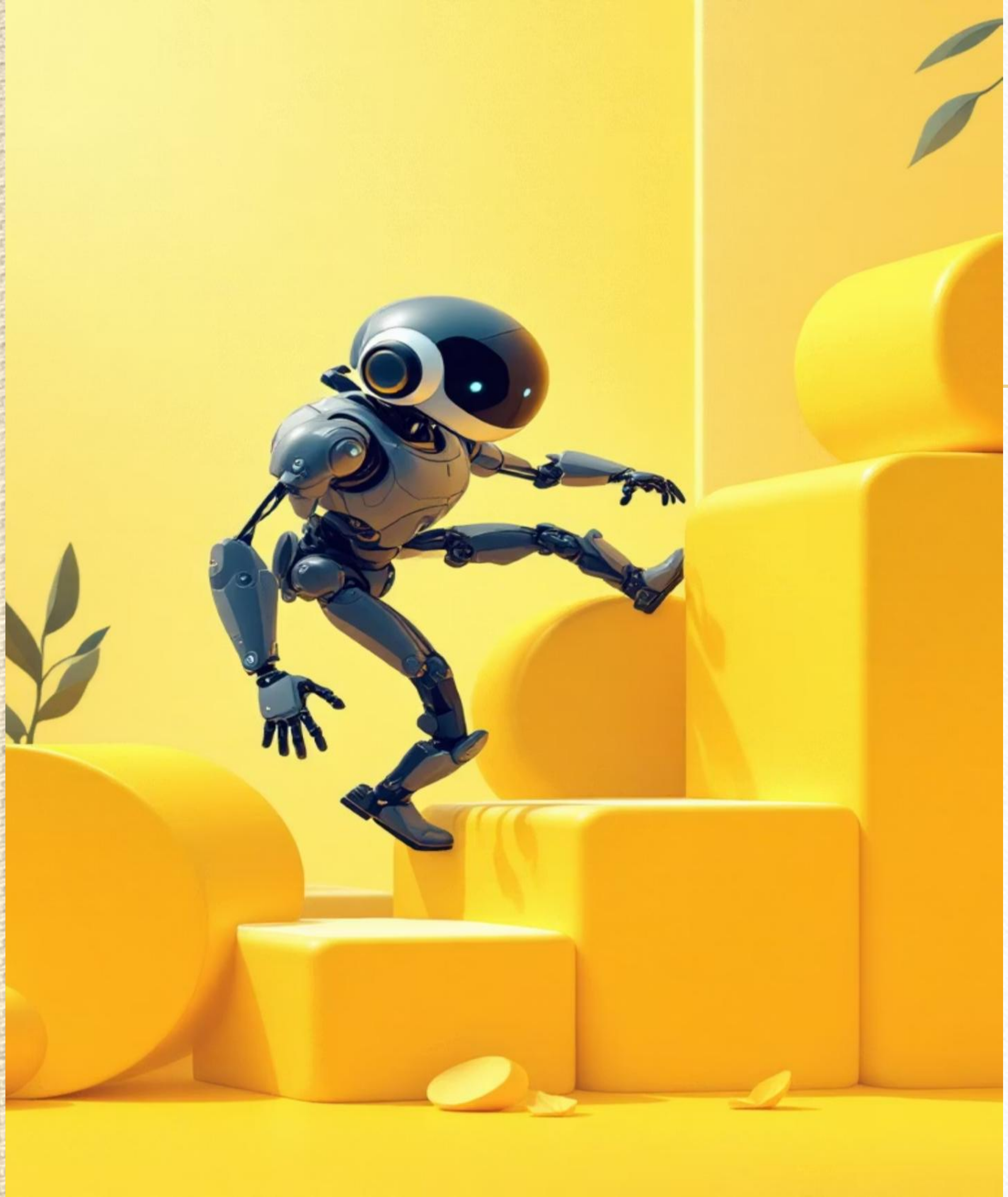
Deployment

XGBoost Model can be integrated into:

- Social Media Monitoring Dashboards.
- Real-time Brand Sentiment Trackers.
- Customer Feedback Systems.

Recommendations

- **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?