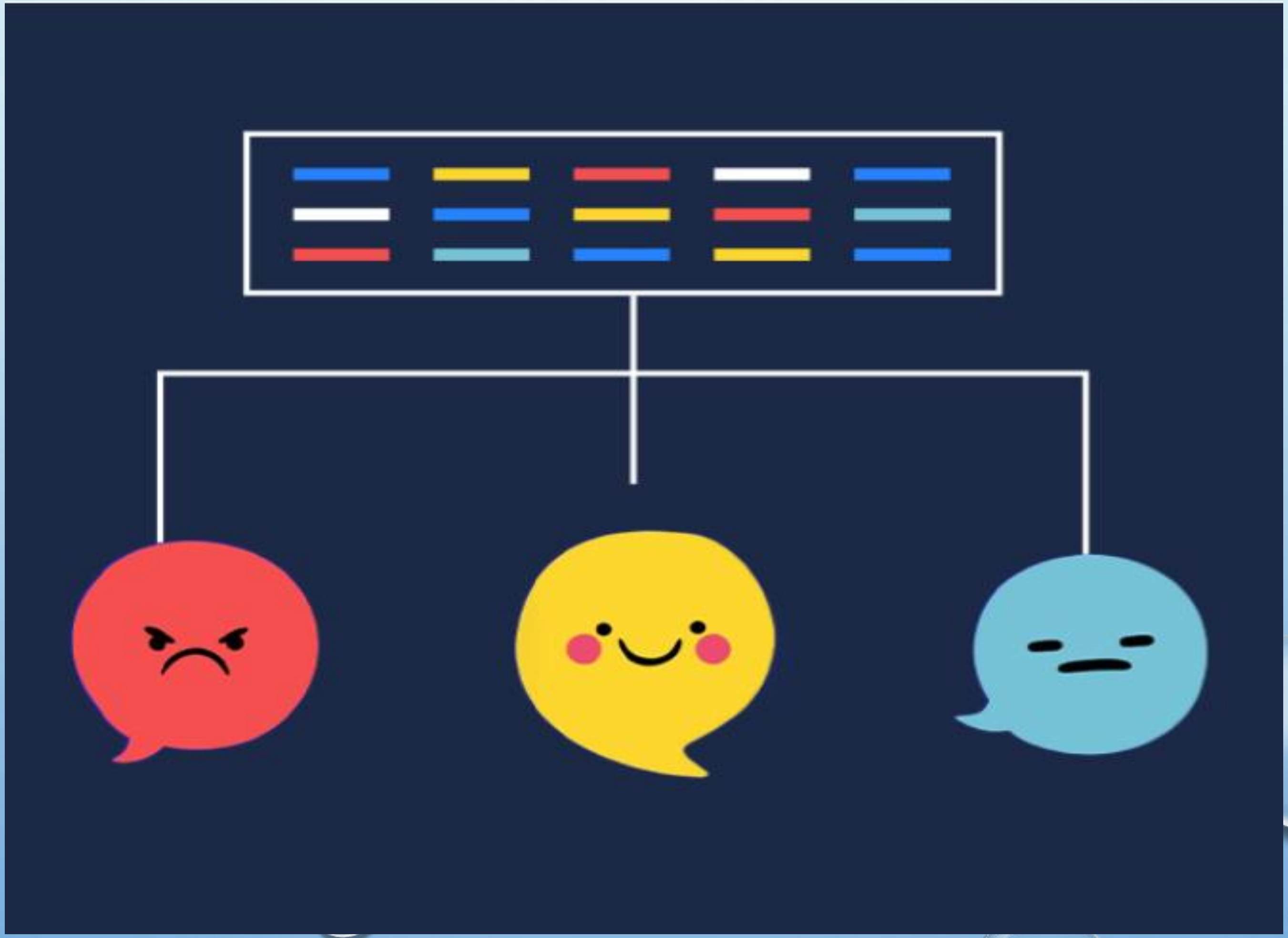
FROM TWEETS TO STRATEGY :
PREDICTING SENTIMENT, SHAPING STRATEGY



BUSINESS UNDERSTANDING

IN THE CURRENT DIGITAL LANDSCAPE, SOCIAL MEDIA PLATFORMS SUCH AS X (FORMERLY TWITTER) HAVE BECOME POWERFUL SPACES WHERE CONSUMERS EXPRESS THEIR OPINIONS, EXPERIENCES, AND PERCEPTIONS ABOUT BRANDS AND PRODUCTS. THESE ONLINE CONVERSATIONS OFFER ORGANIZATIONS A VALUABLE OPPORTUNITY TO MEASURE CUSTOMER SENTIMENT, EVALUATE BRAND PERCEPTION, AND SHAPE STRATEGIES THAT DRIVE COMPETITIVE ADVANTAGE.

BUSINESS GOALS

- 1. ANALYZE CUSTOMER PERCEPTION OF APPLE AND GOOGLE BRANDS, INCLUDING THEIR PRODUCTS AND SERVICES.
- 2. BUILD A PREDICTIVE MODEL THAT CLASSIFIES SENTIMENTS OF A TWEET.
- 3. PROVIDE ACTIONABLE INSIGHTS THAT HELP SHAPE MARKETING CAMPAIGNS, IMPROVE CUSTOMER SERVICE, AND STRENGTHEN ENGAGEMENT STRATEGIES.

TARGET AUDIENCE

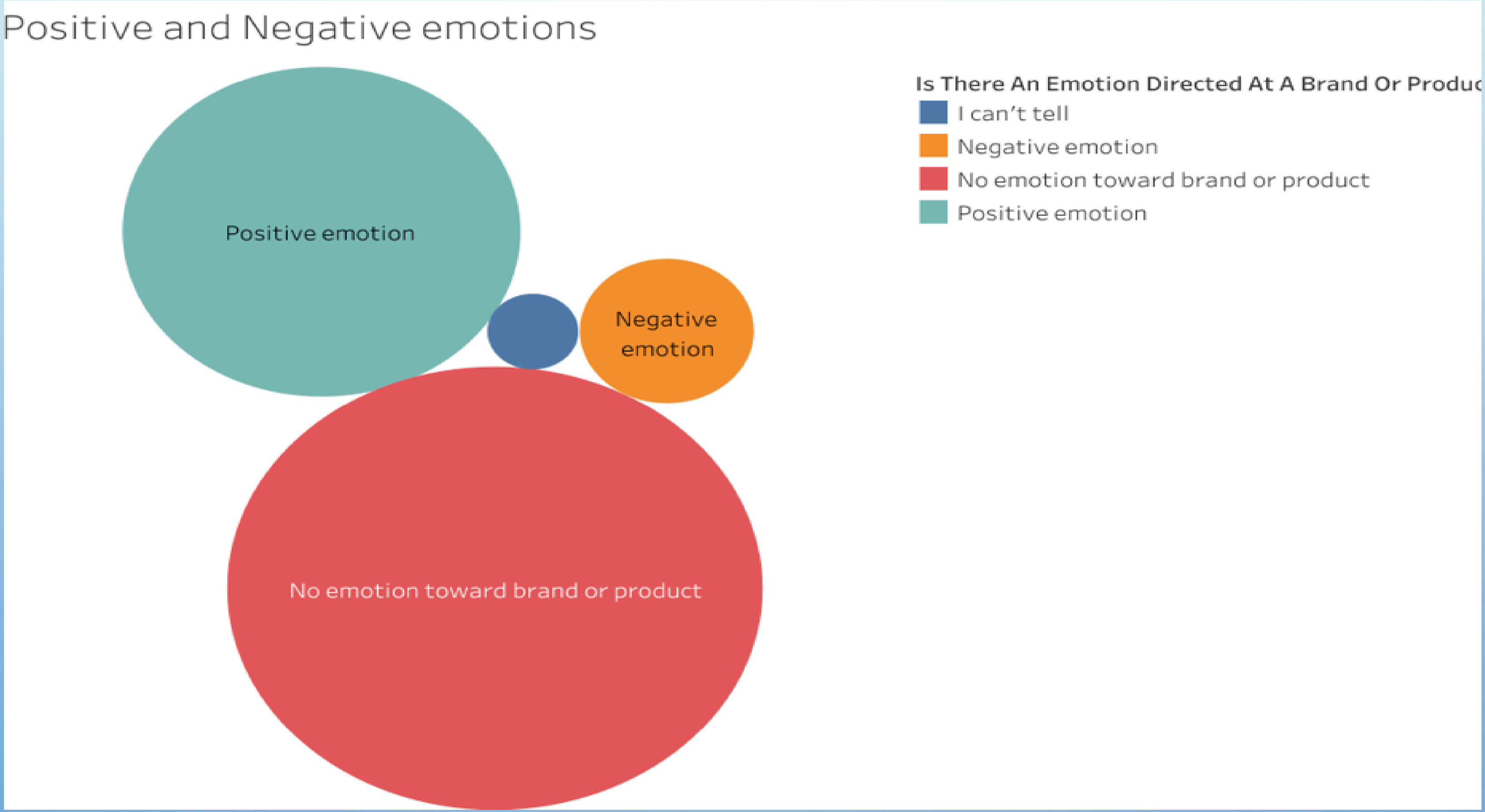
- MARKETING TEAMS AT APPLE AND GOOGLE
- CUSTOMER ENGAGEMENT & EXPERIENCE MANAGERS.
- PRODUCT MANAGERS (APPLE & GOOGLE DIVISIONS SUCH AS IPHONE, PIXEL, ANDROID, IOS)
- COMPETITIVE STRATEGY ANALYSTS.
- BUSINESS EXECUTIVES & SENIOR LEADERSHIP

DATA UNDERSTANDING

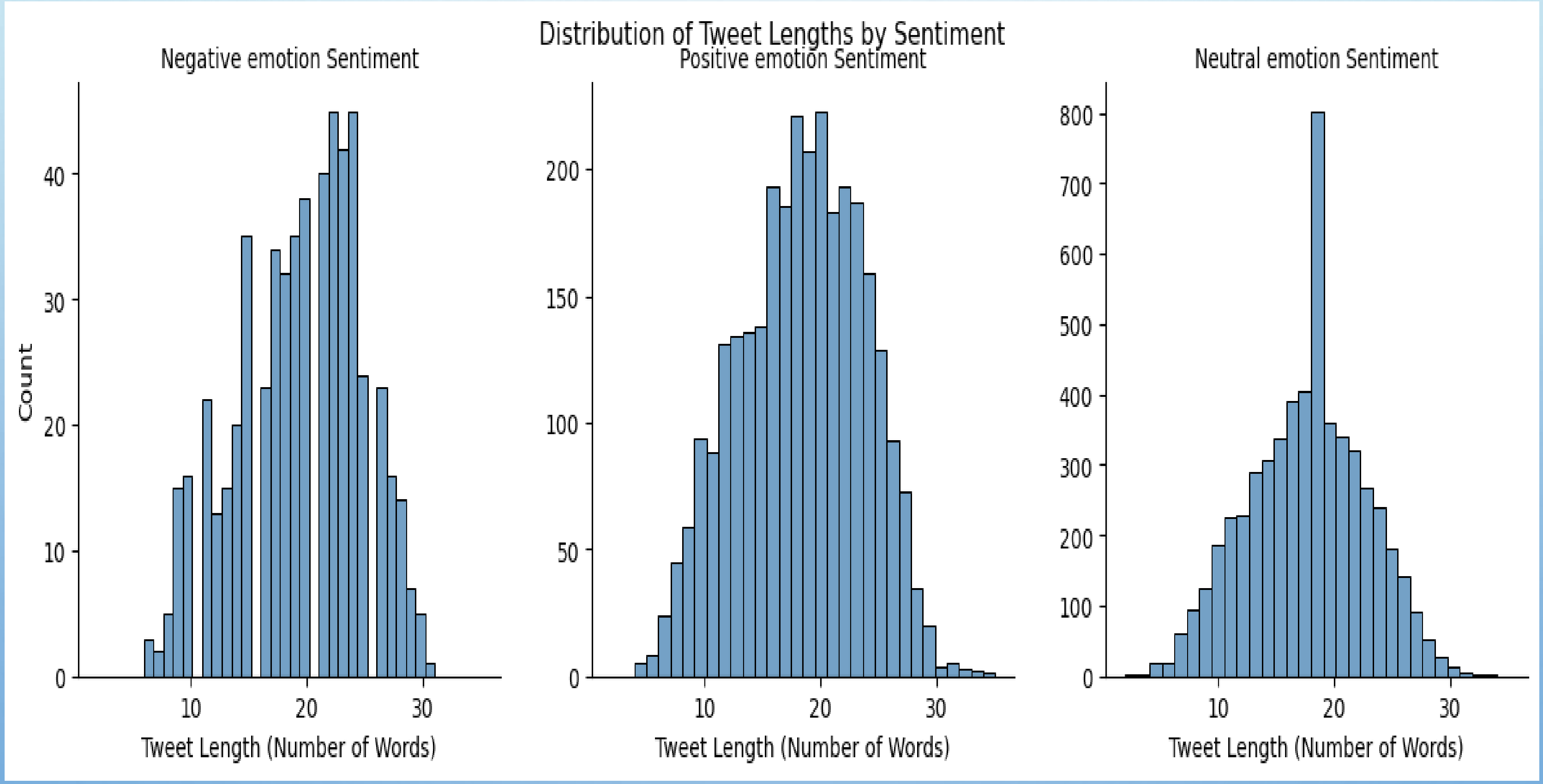
- DATA CLEANING: REMOVE NOISE (URLS, MENTIONS), LOWERCASE, TOKENIZE
- FEATURE EXTRACTION: BAG-OF-WORDS / TF–IDF
- BUILD A PREDICTIVE MODEL THAT CLASSIFIES SENTIMENTS OF A TWEET.
- MODELING: BASELINE CLASSIFIERS AND EVALUATION (ACCURACY, F1)

DATA SOURCE: DATA.WORLD

PLOTTING TWEETS SENTIMENT DISTRIBUTION



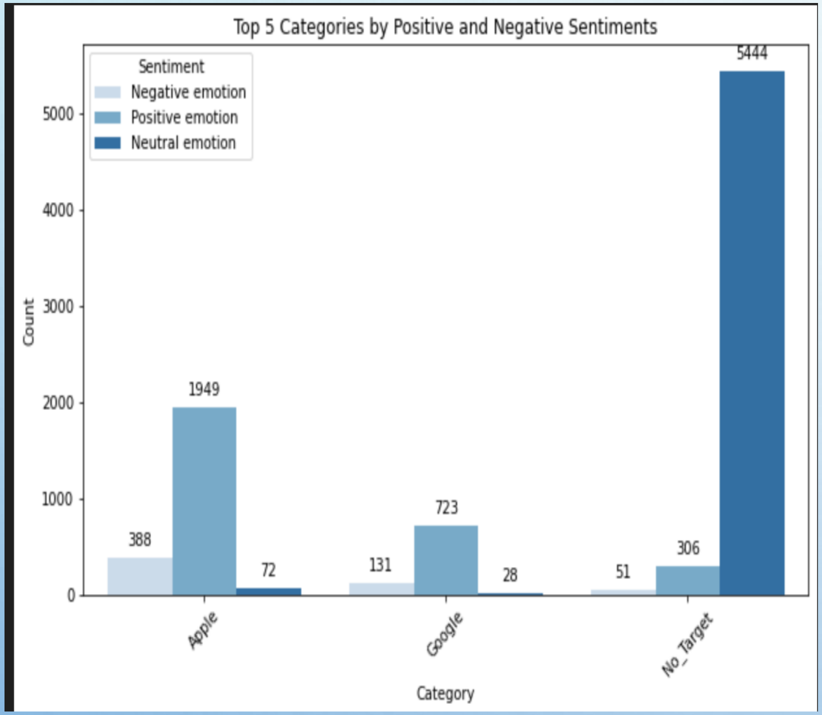
DISTRIBUTION OF TWEET LENGTH BY SENTIMENT

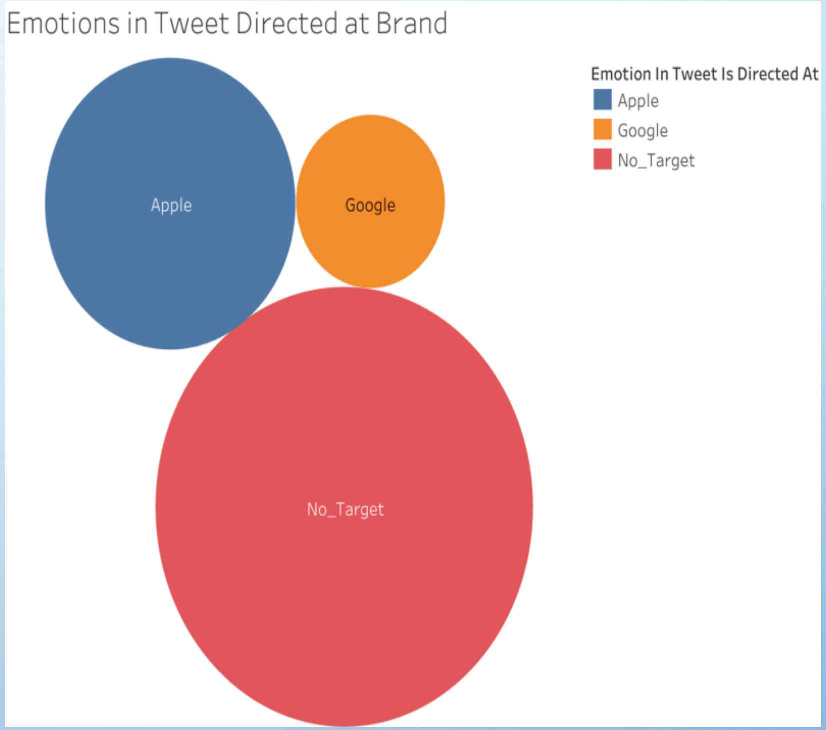


The sentiment distribution indicates that the majority of conversations are neutral, with 5,544 tweets, representing approximately 60.6% of the dataset. This suggests that while consumers are actively discussing brands and products, a large portion of the discourse remains informational or unbiased, lacking strong emotional tones.

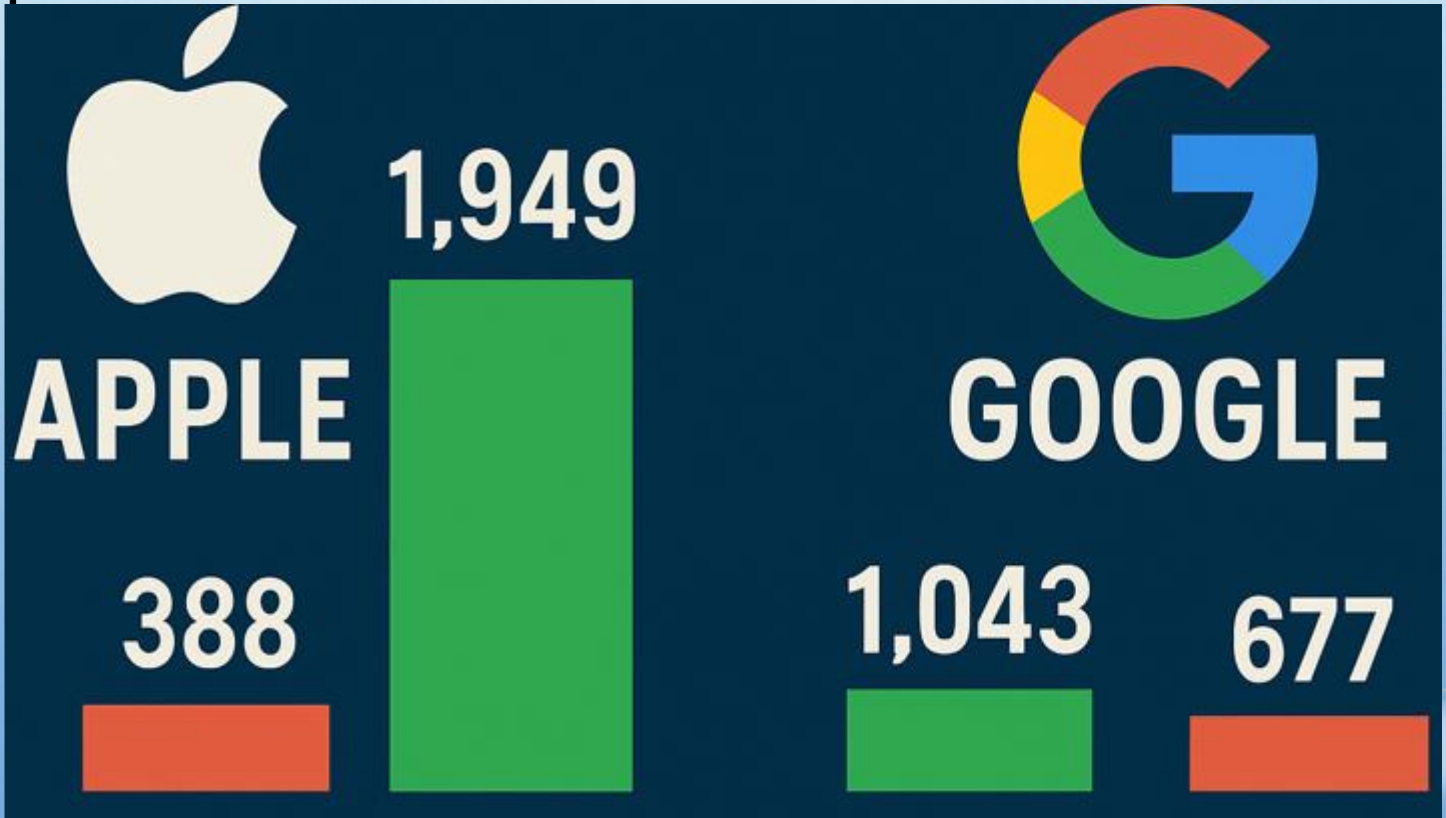
Positive sentiment follows with 2,978 tweets (32.6%), highlighting a considerable base of satisfied or supportive consumers who express favorable opinions. Meanwhile, negative sentiment is minimal, at only 570 tweets (6.2%), indicating relatively low levels of dissatisfaction or criticism compared to the overall conversation volume.

BRAND SENTIMENTS ANALYSIS





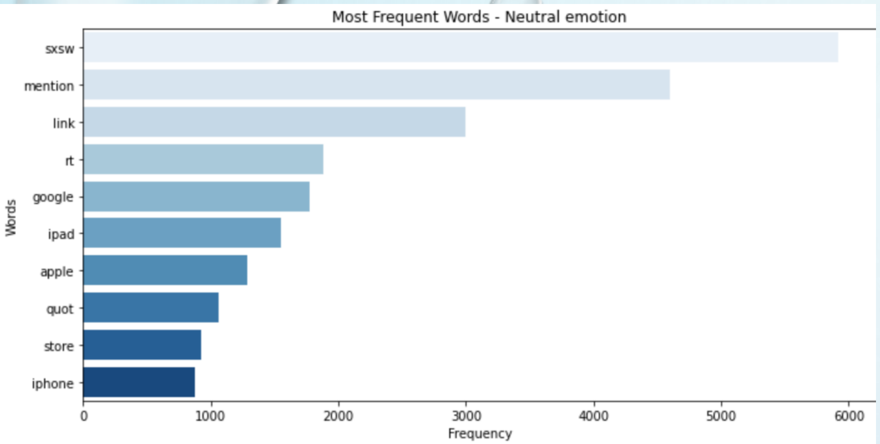
Apple holds a stronger brand perception compared to Google within the dataset from the analysis. Apple products and the brand generate the highest level of engagement, with 1,949 positive mentions against 388 negative mentions, indicating strong consumer affinity despite some criticism.

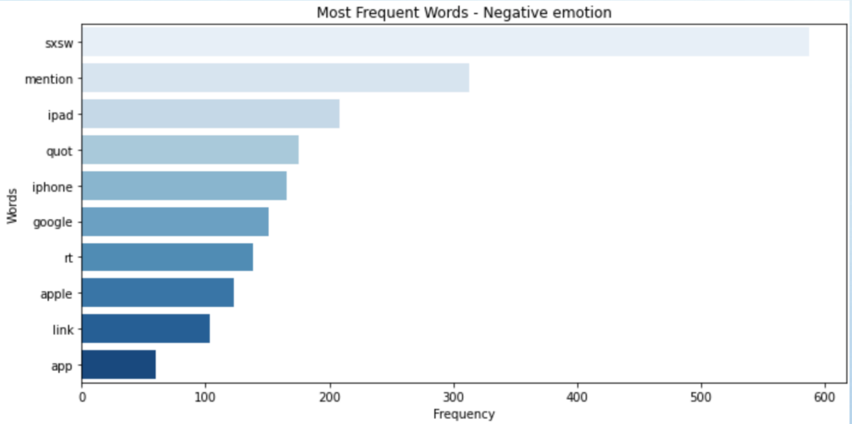


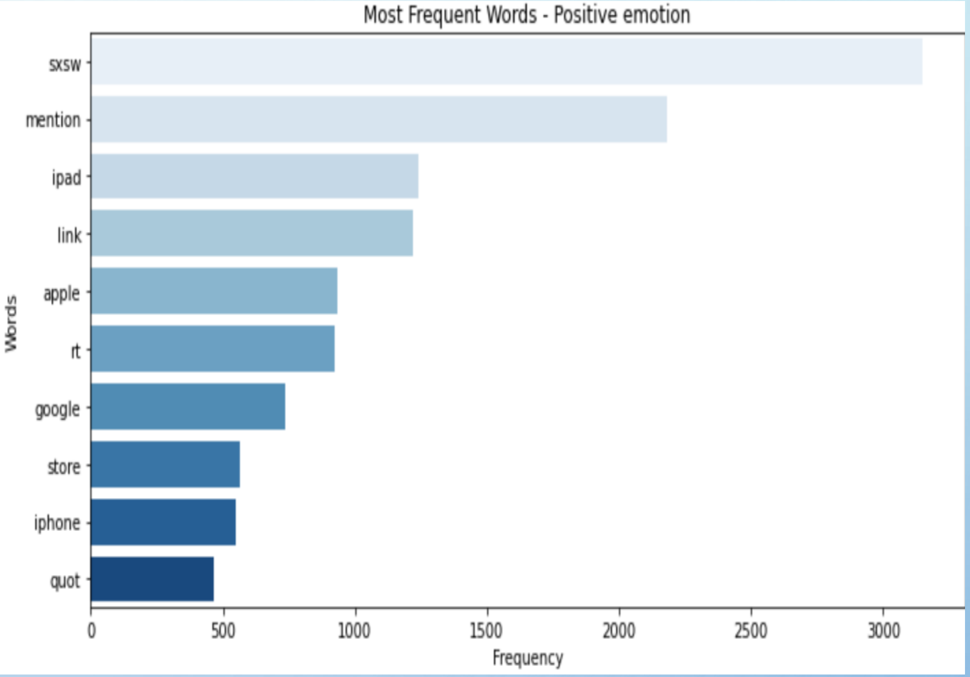
- Negative tweets has a clear peak around 22–24 words. Very short tweets (<10 words) and very long tweets (>30 words) are rare. The distribution is right-skewed, suggesting that while most negative tweets are moderately long, a few extend further in length.

- Positive tweets curve peaks around 18–20 words, ranging between 10 and 26 words. The distribution is also right-skewed, indicating that users tend to use slightly more words when expressing positive sentiment.

- Neutral tweets follow a normal distribution with a strong central peak at 18 words, showing a consistent word count pattern , suggesting that neutral tweets are more standardized, possibly because they are more factual or informational.


Most Frequent Words - Neutral emotion
Words
sxsw
mention
link
rt
google
ipad
apple
quot
store
iphone
0
1000
2000
3000
4000
5000
6000
Frequency


Most Frequent Words - Negative emotion
Words
sxsw
mention
ipad
quot
iphone
google
rt
apple
link
app
0
100
200
300
400
500
600
Frequency


Most Frequent Words - Positive emotion
Words
sxsw
mention
ipad
link
apple
rt
google
store
iphone
quot
0
500
1000
1500
2000
2500
3000
Frequency

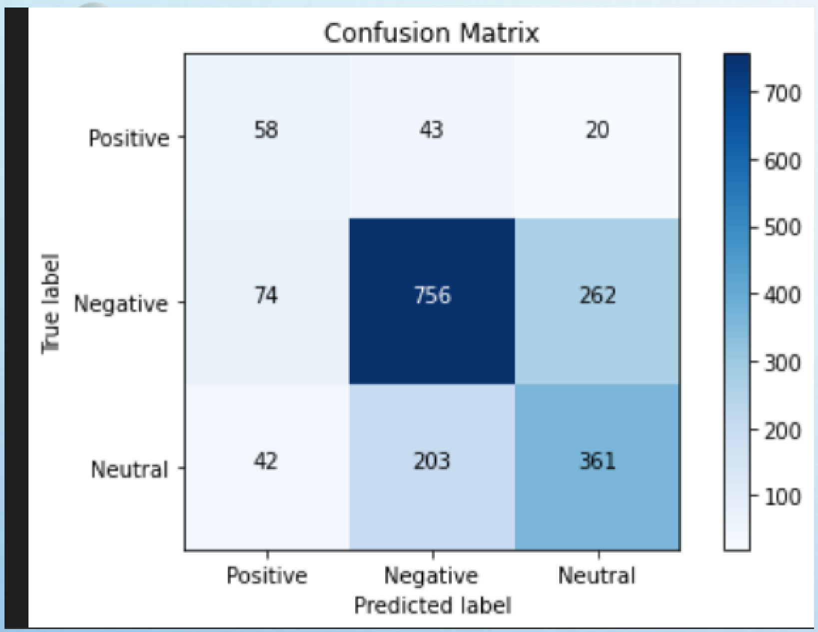
Discussions on Twitter were largely event-driven, with “sxsw” dominating across negative, positive, and neutral tweets.

- Negative tweets focused on dissatisfaction with the event or frustrations with tech products like the iPad, iPhone, Apple, and Google.

- Positive tweets reflected enthusiasm and praise for the same event and products, underscoring how shared experiences can generate both excitement and criticism.

- Neutral tweets were largely informational, consisting of news, updates, and links about SXSW and tech launches, without strong emotional tones.

MODELLING



Logistic Regression was used due to its superior precision in classifying tweet sentiments. This makes it especially effective for identifying positive sentiment with minimal false positives, aligning well with our goal of extracting reliable brand perception insights.

EDA revealed a significant imbalance in sentiment classes, with *neutral* tweets dominating the dataset and *negative* sentiment underrepresented. SMOTE was used to address during model training using Logistic Regression. The model automatically adjusts the importance of each class based on its frequency, penalizing misclassification of minority classes more heavily. This approach helps improve recall and F1-score for underrepresented sentiments without altering the original data distribution.

MODEL PERFROMANCE

- ACCURACY: 0.647

RECOMMENDATIONS

1. Leverage Neutral Sentiment for Engagement - Since most tweets are neutral, brands should actively convert these mentions into advocacy through personalized responses, storytelling, and targeted campaigns.

2. Improve Detection of Negative Sentiment - Current models underperform on negative tweets. Prioritize real-time monitoring of negative sentiment to address complaints quickly and protect brand reputation.

3. Brand-Specific Strategies -
- Apple: Strong visibility but polarized opinions. Focus on addressing customer pain points while amplifying advocacy from loyal users.
- Google: Stable but lower engagement. Invest in campaigns that generate excitement to match Apple's market buzz.

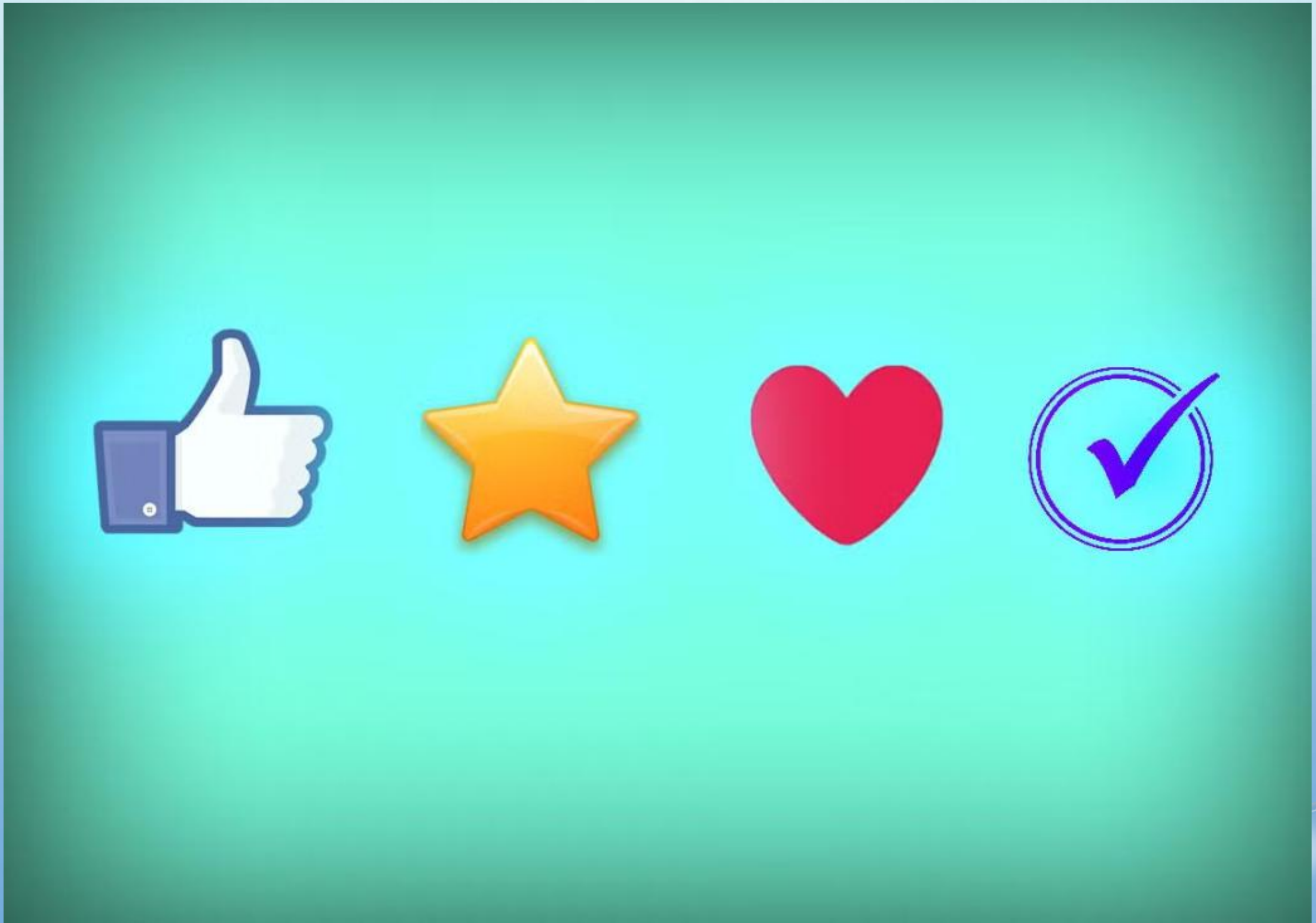
4. Event-Driven Monitoring - Spikes in sentiment were tied to major events (e.g., SXSW). Companies should implement real-time sentiment dashboards to capitalize on product launches, events, and trending topics.

5. Class Imbalance Handling - Continue using SMOTE and resampling techniques to improve recall for minority classes (especially negative sentiment). Consider cost-sensitive learning or ensemble methods to further balance performance.

WAY FORWARD

- - NEUTRAL TWEETS DOMINATE, REPRESENTING UNTAPPED OPPORTUNITIES FOR ENGAGEMENT.
- - POSITIVE SENTIMENT IS CAPTURED MODERATELY WELL, REFLECTING BRAND LOYALTY AND CONSUMER ADVOCACY.
- - NEGATIVE SENTIMENT DETECTION REMAINS A CHALLENGE, REQUIRING ADVANCED NLP TECHNIQUES FOR IMPROVEMENT.

Q&A



THANK YOU

PROJECT COLLABORATORS

- AGNES CHOMBA
- OLGAH OMOLLO
- DERRICK MALING'A
- ERIC OKACHA
- JUDAH ODIDAH
- NICK MWAI
- LUCAS OMINDE