Group7

DECODING BRAND SENTIMENTS: A MACHINE LEARNING APPROACH

Presented by Group Seven:

Doreen Chepkong'a, Boniface Thuo, Linda Temoet
Grace Mwega, Micheal Munene

CONTENT

1. **Introduction**
   - Background Information
   - Challenges in Sentiment Analysis
   - Stakeholders
   - Proposed Solution (Analysis & Modeling)
   - Objective
   - Metrics of Success
2. **Data Understanding**
3. **Data reparation & Analysis**
4. **Modeling**
5. **Evaluation**
6. **Conclusion**
7. **Recommendation**

INTRODUCTION

Background Information

In today's digital era, Twitter and other social media platforms profoundly influence consumer opinions about brands and products. Tweets serve as a reflection of public sentiment—positive, negative, or neutral—toward specific brands or products. These sentiments significantly shape brand reputation, consumer trust, and buying decisions.

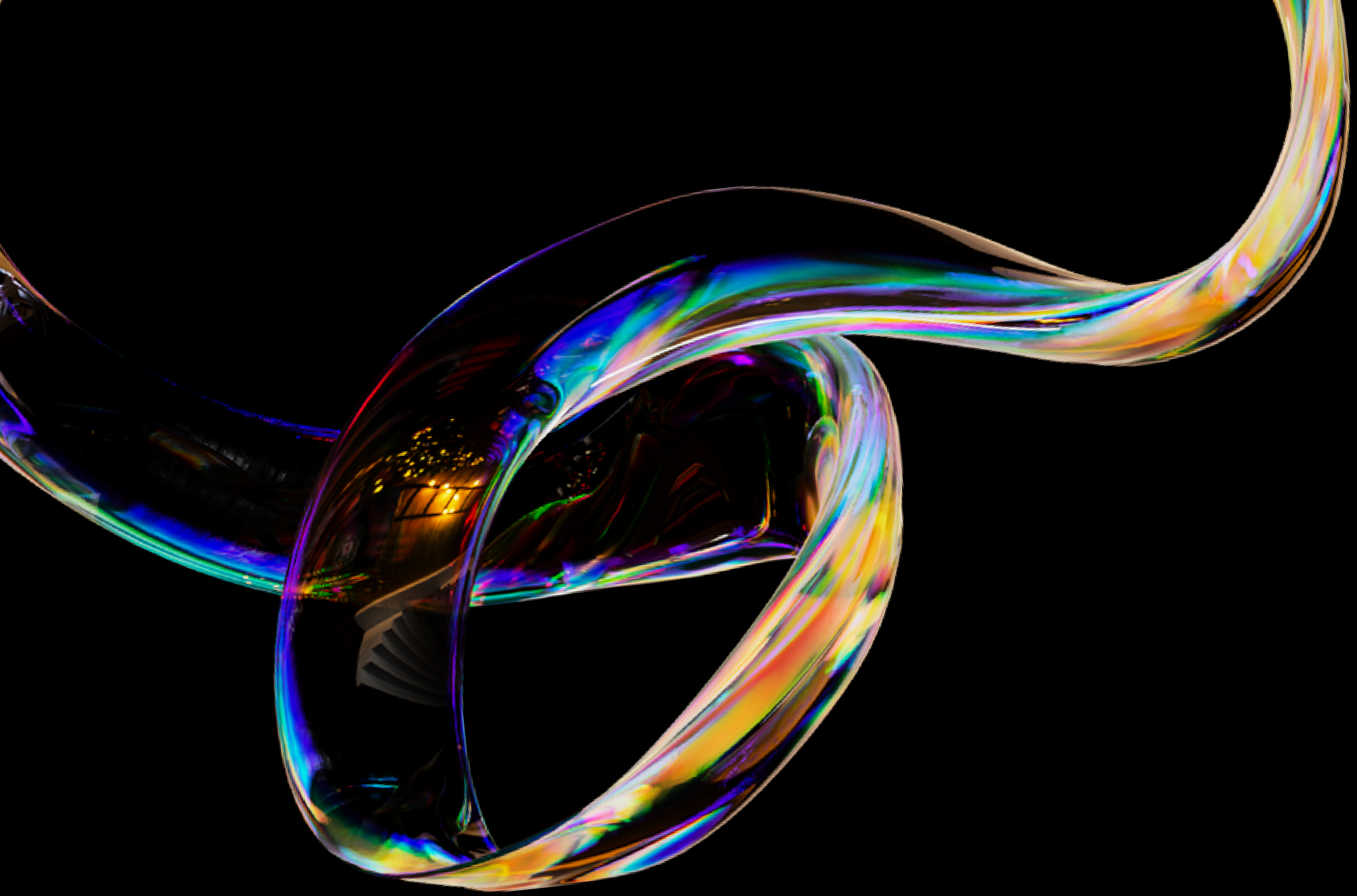
Challenges in Sentiment Analysis

- The overwhelming volume of tweets makes manual analysis impractical and time-consuming.
- Tweets often contain slang, emojis, and abbreviations, which complicate text analysis.
- Accurately distinguishing between similar sentiments (e.g., sarcasm vs. genuine praise) poses a challenge.

PROPOSED SOLUTION

(Analysis & Modeling)

The project proposes a machine learning-based sentiment analysis system using Natural Language Processing (NLP) techniques. The model will classify tweets about Apple and Google products into sentiment categories (positive, neutral, negative), providing actionable insights for marketing, product development, and competitive benchmarking



OBJECTIVE

The aim of this project is to design and implement a sentiment analysis system that accurately classifies tweets into various sentiment categories, providing actionable insights for brands and product stakeholders.

- Develop an NLP model to classify tweets about Apple and Google products as positive, negative, or neutral, providing actionable insights into customer sentiment.

- Optimize the model's performance using text preprocessing, feature engineering, and iterative evaluations to ensure high accuracy and reliability.

- Analyze sentiment trends to support strategic decision-making for marketing, product development, and competitive benchmarking

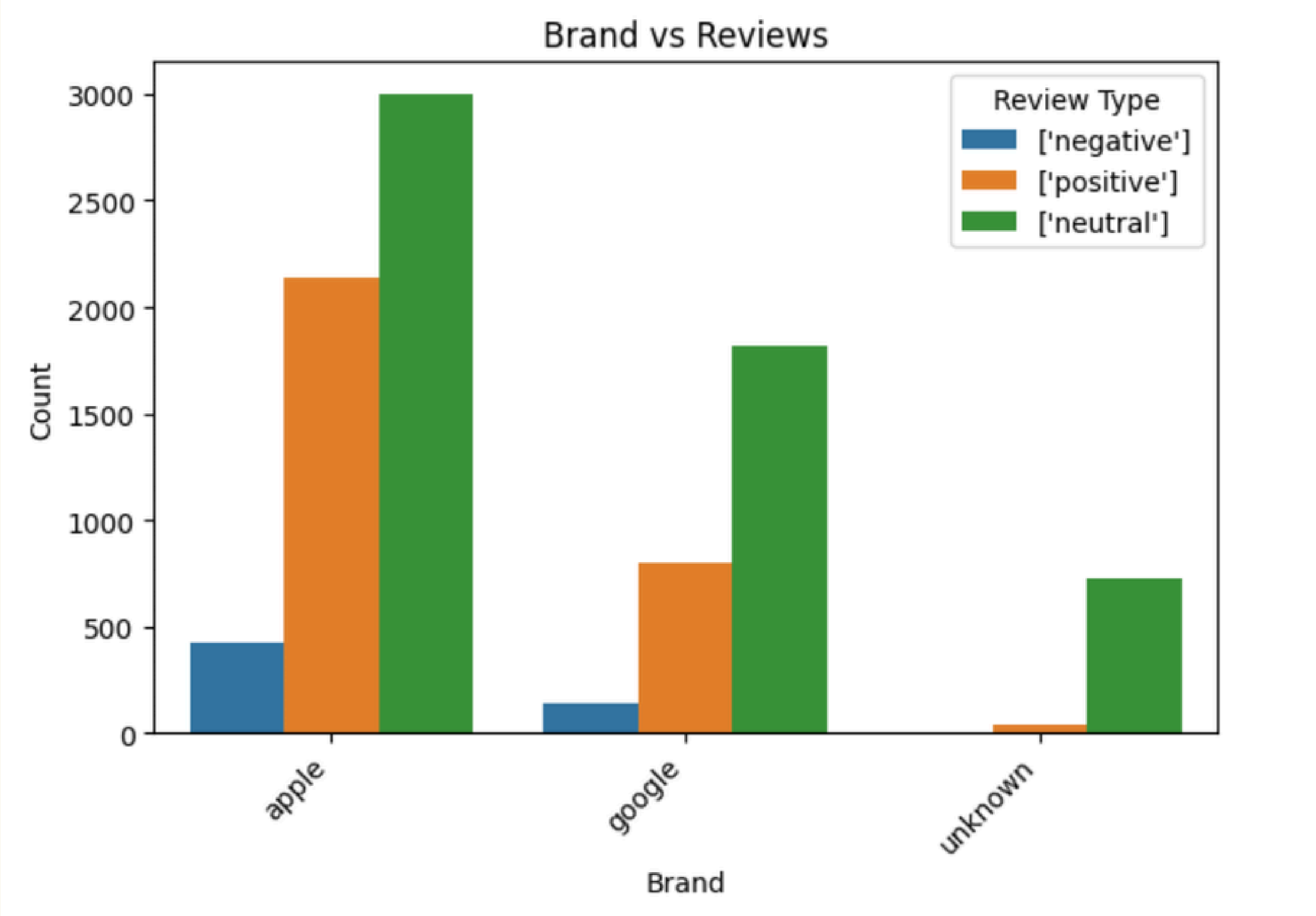
- Compare the two brands' perceptions by people by analyzing which brand has more positive tweets and negative tweets.

DATA UNDERSTANDING

The data for this analysis is extracted from CrowdFlower via data.world. The dataset contains over 9,000 tweets that reflect emotional reactions to brands or products. It includes columns for the text of the tweet, the targeted brand, product, or event, and the sentiment direction (positive or negative). It contains three columns that holds the following meaning:

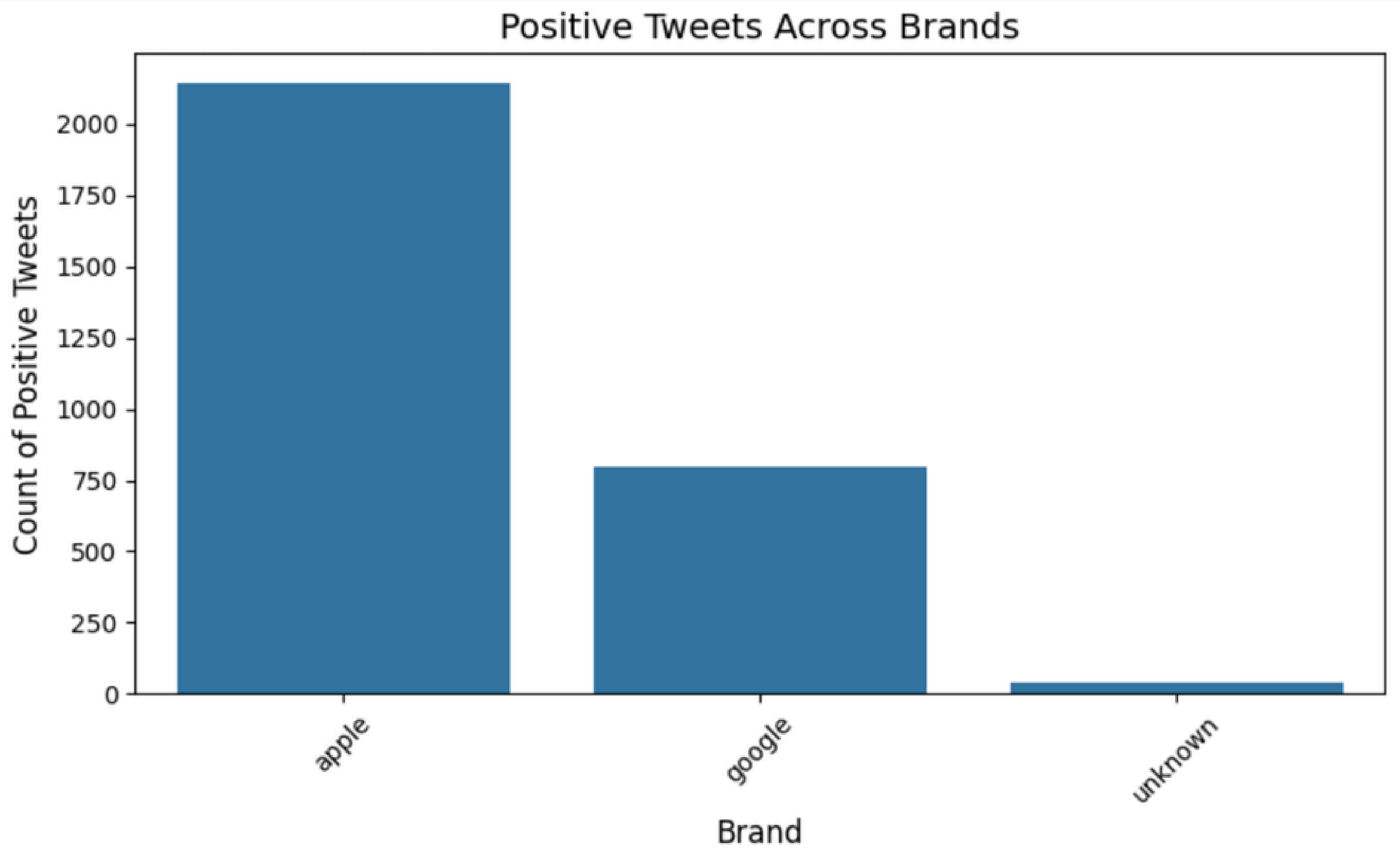
- emotion in tweet is directed at: Indicates the entity or topic the emotion in the tweet is directed towards.
- tweet text: Contains the actual text of the tweet.
- is_there_an_emotion_directed_at_a_brand_or_product: The target variable, indicating whether an emotion is expressed towards a brand or product.

BRAND VERSUS REVIEWS



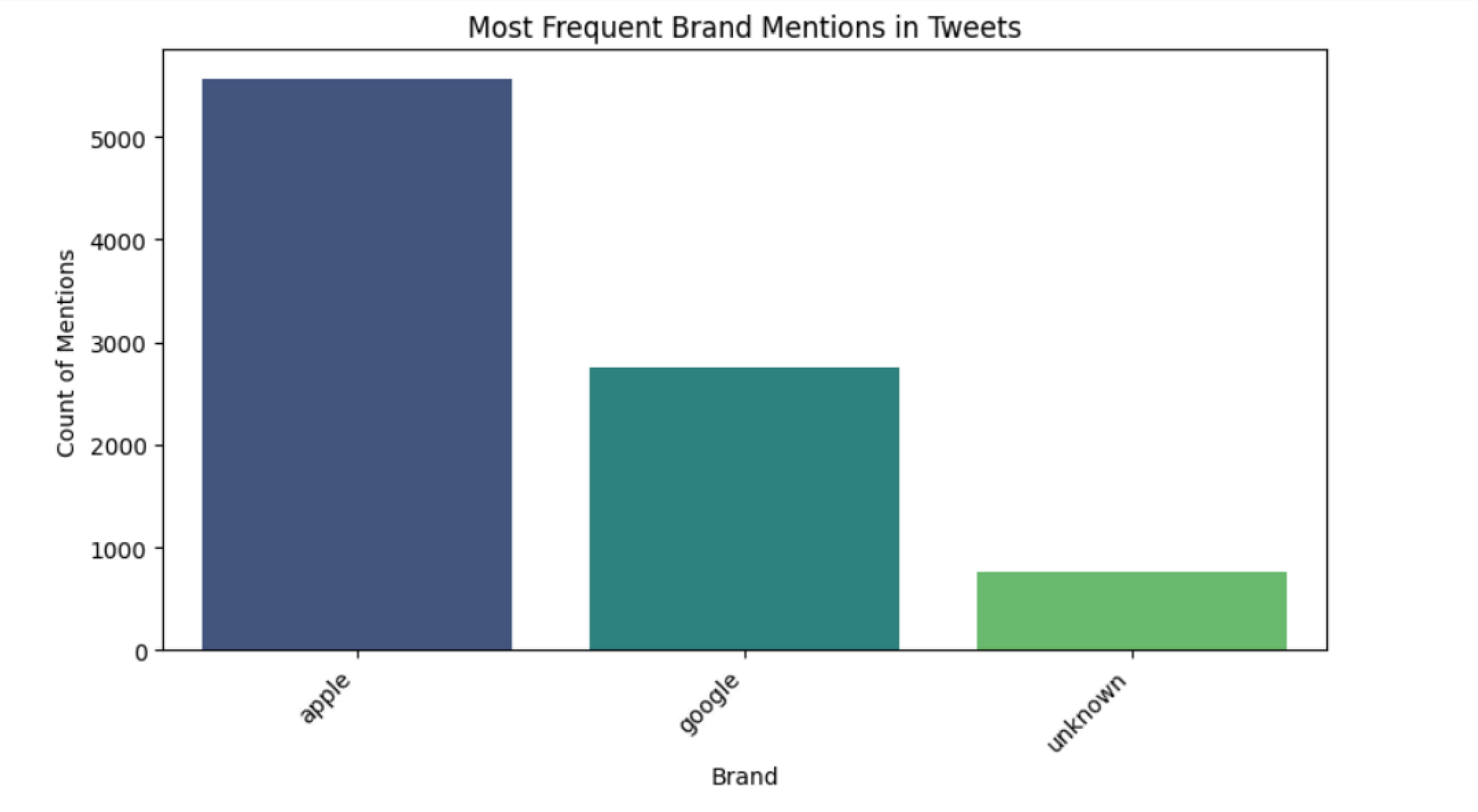
The graph above shows the distribution of positive, negative, and neutral reviews across brands: Apple, Google, and unknown. Apple has the highest positive and neutral reviews, while Google has fewer overall reviews. Neutral reviews dominate across all brands, with Apple having the most significant share.

POSITIVE COUNTS FOR ALL BRANDS.



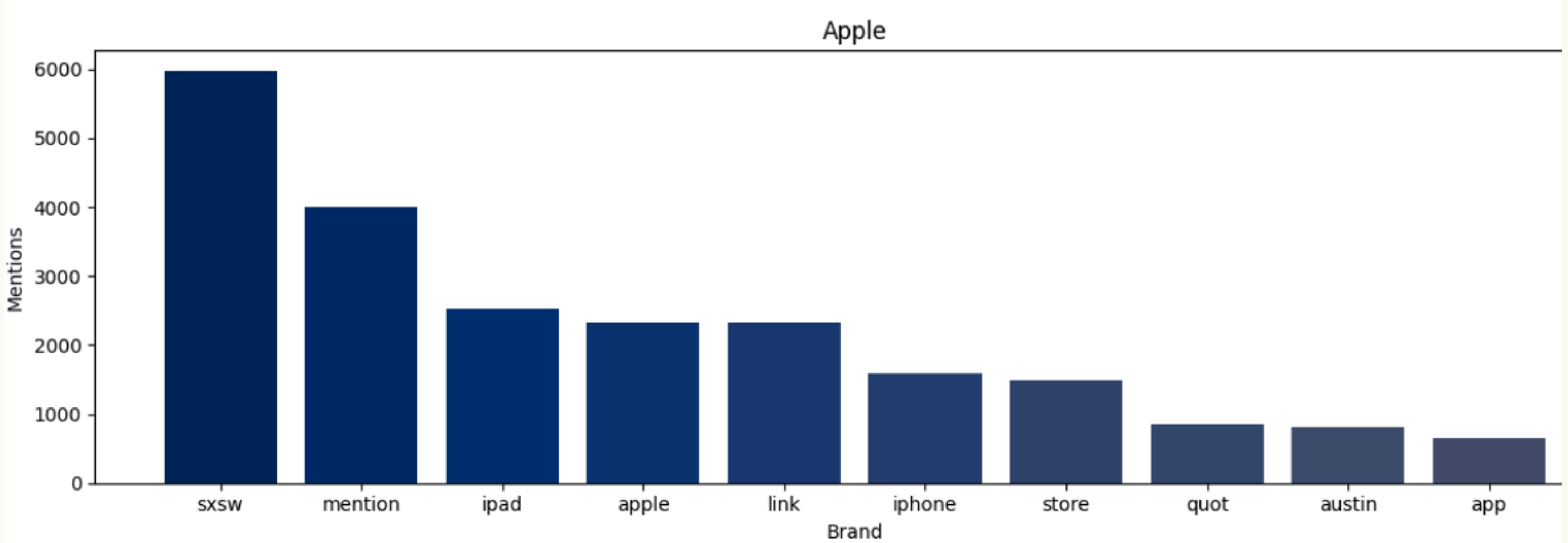
The graph illustrates the distribution of positive tweets directed at three brands: Apple, Google, and Unknown. Apple leads with the highest number of positive tweets, followed by Google. The "Unknown" category has the least. This suggests that Apple garners more positive sentiment compared to the other brands in the dataset.

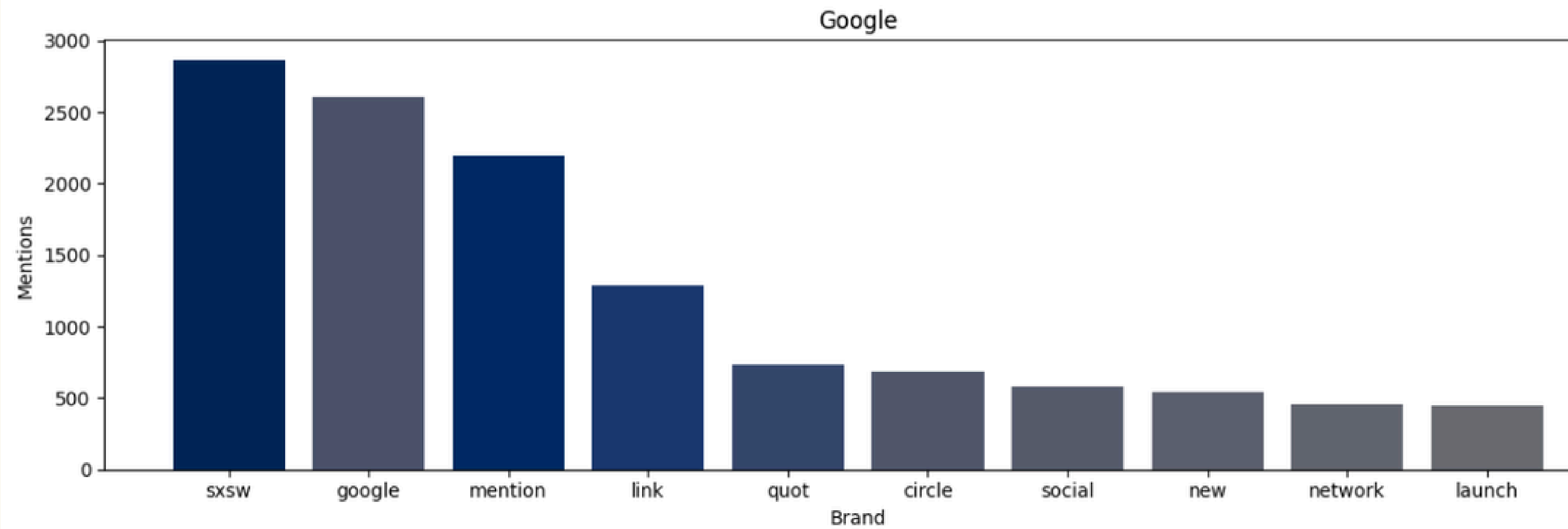
MOST FREQUENT BRAND MENTIONS IN TWEETS

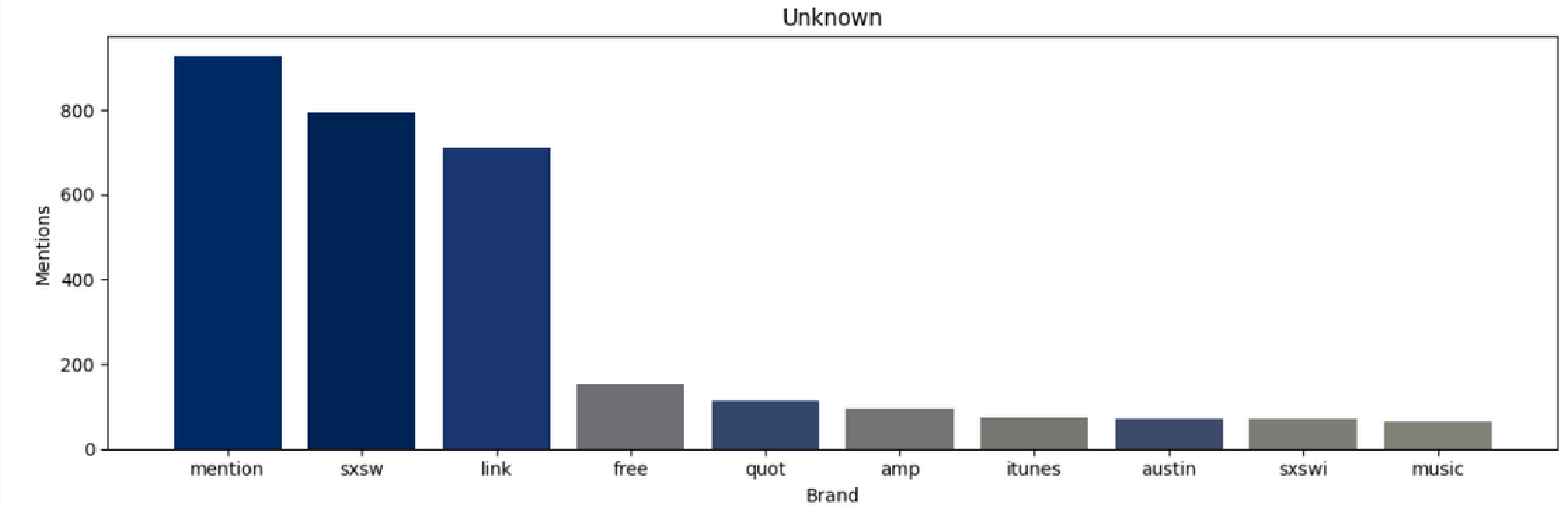


This graph above shows the count of brand mentions in tweets. Apple has the highest mentions, followed by Google, while "unknown" brands have significantly fewer mentions. The visualization highlights the dominance of Apple in the dataset.

BRAND VERSUS REVIEWS

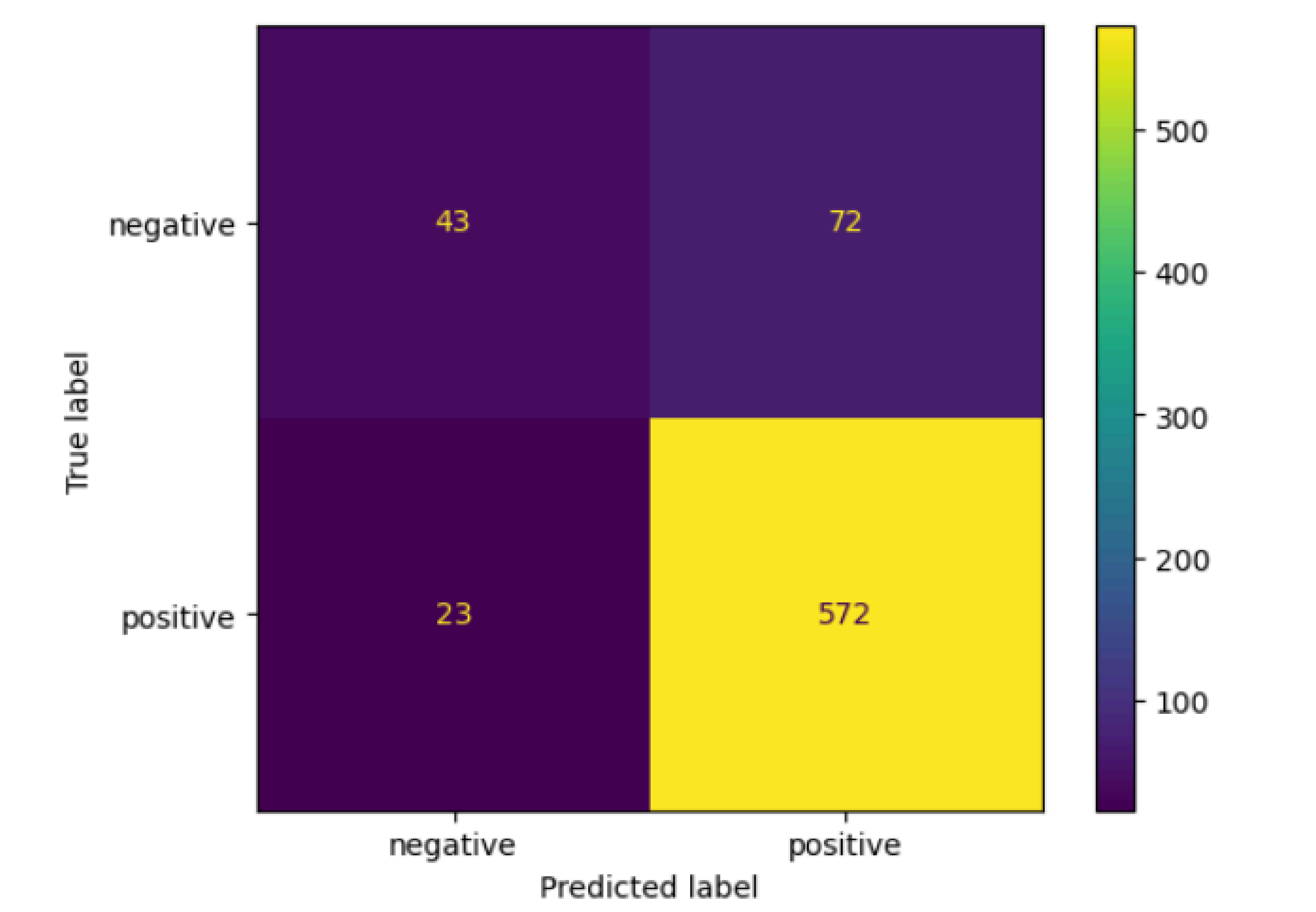






These graphs show the most frequent words that have been used for all the brands. With "sxsw" being the most frequent for Apple and Google and "mention" being the most frequent for the other brands.

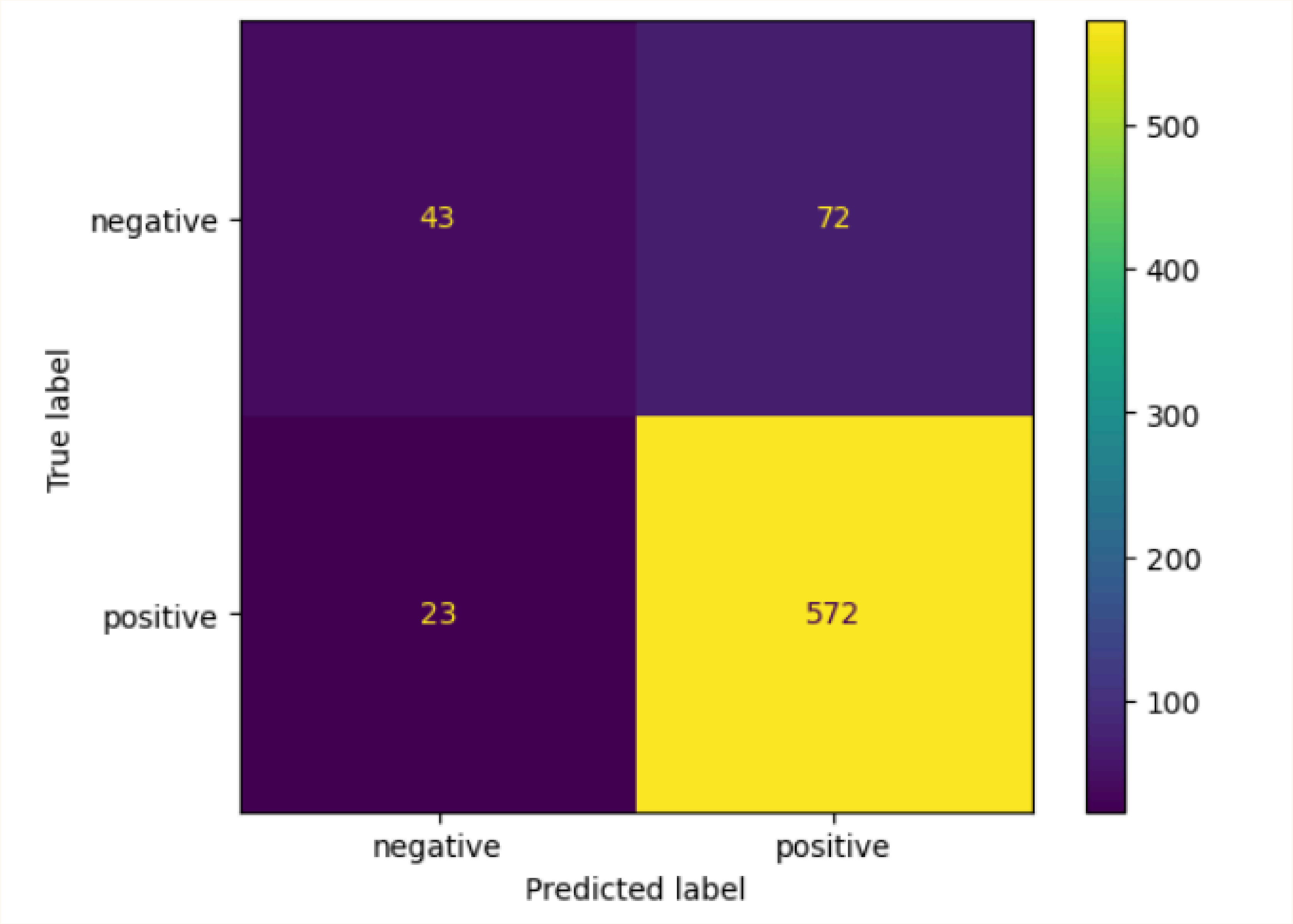
Logistic Regression



The confusion matrix shows strong model performance, with high true positive and true negative values. Accuracy is 88.9%, indicating the percentage of correct predictions. Precision is 87.8%, reflecting prediction reliability. Recall is 88.9%, measuring sensitivity. The F1 score of 87.4% balances precision and recall, confirming robust classification performance.

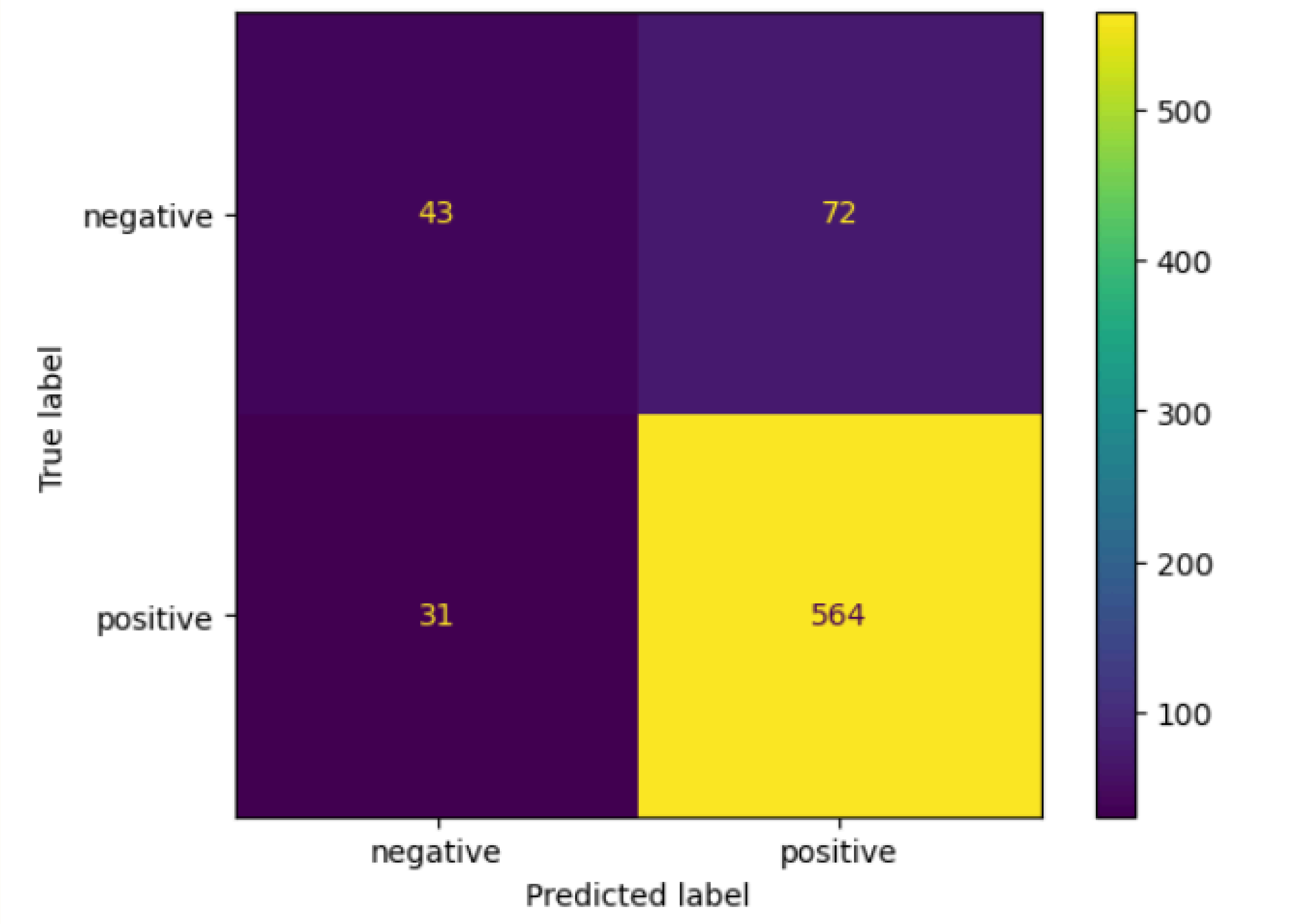
In summary, the Logistic Regression model demonstrates strong performance in classifying the sentiment of tweets, with high accuracy, precision, recall, and F1-score. This indicates that the model is reliable in identifying positive tweets and has a low rate of false positives and false negatives.

Weighted Logistic Regression



These metrics suggest that the model is performing well. It has a high accuracy, indicating overall correctness. It also has good precision and recall, meaning it is both reliable in its positive predictions and able to identify most of the actual positive cases. The high F1 Score further confirms the model's strong performance.

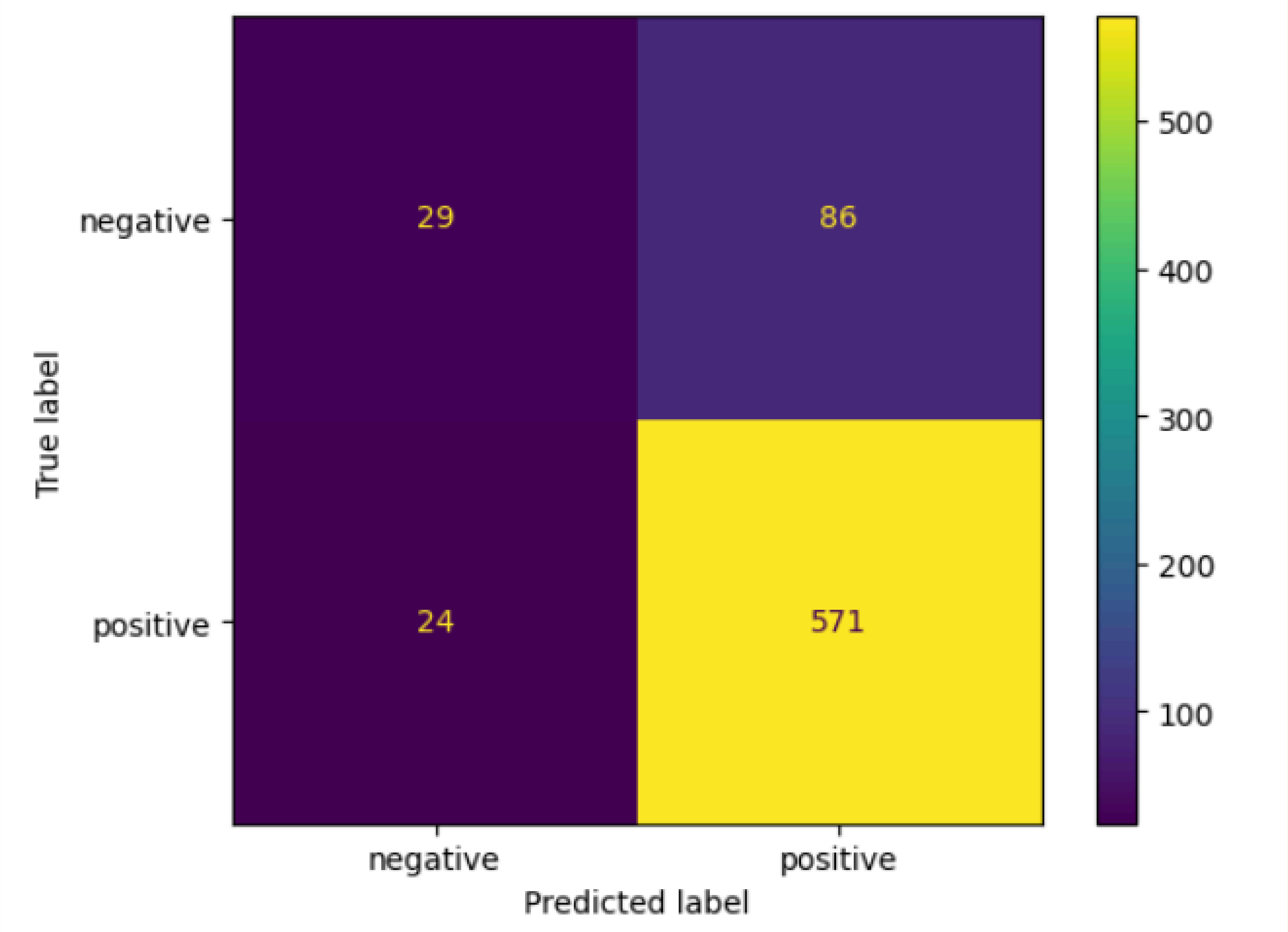
Decision Tree Classifier



The confusion matrix indicates 540 true positives and 46 true negatives, reflecting strong classification performance. There are 35 false negatives and 55 false positives. The accuracy (86.7%) shows overall reliability, while precision (85.7%) highlights the model's positive prediction accuracy. Recall (86.7%) demonstrates sensitivity, and the F1 score (86.1%) confirms a balanced trade-off between precision and recall.

The metrics and confusion matrix indicate that the Decision Tree model is performing well in classifying the sentiment of tweets. It achieves a good balance between precision and recall, suggesting that it can reliably identify both positive and negative sentiments with a low rate of misclassifications.

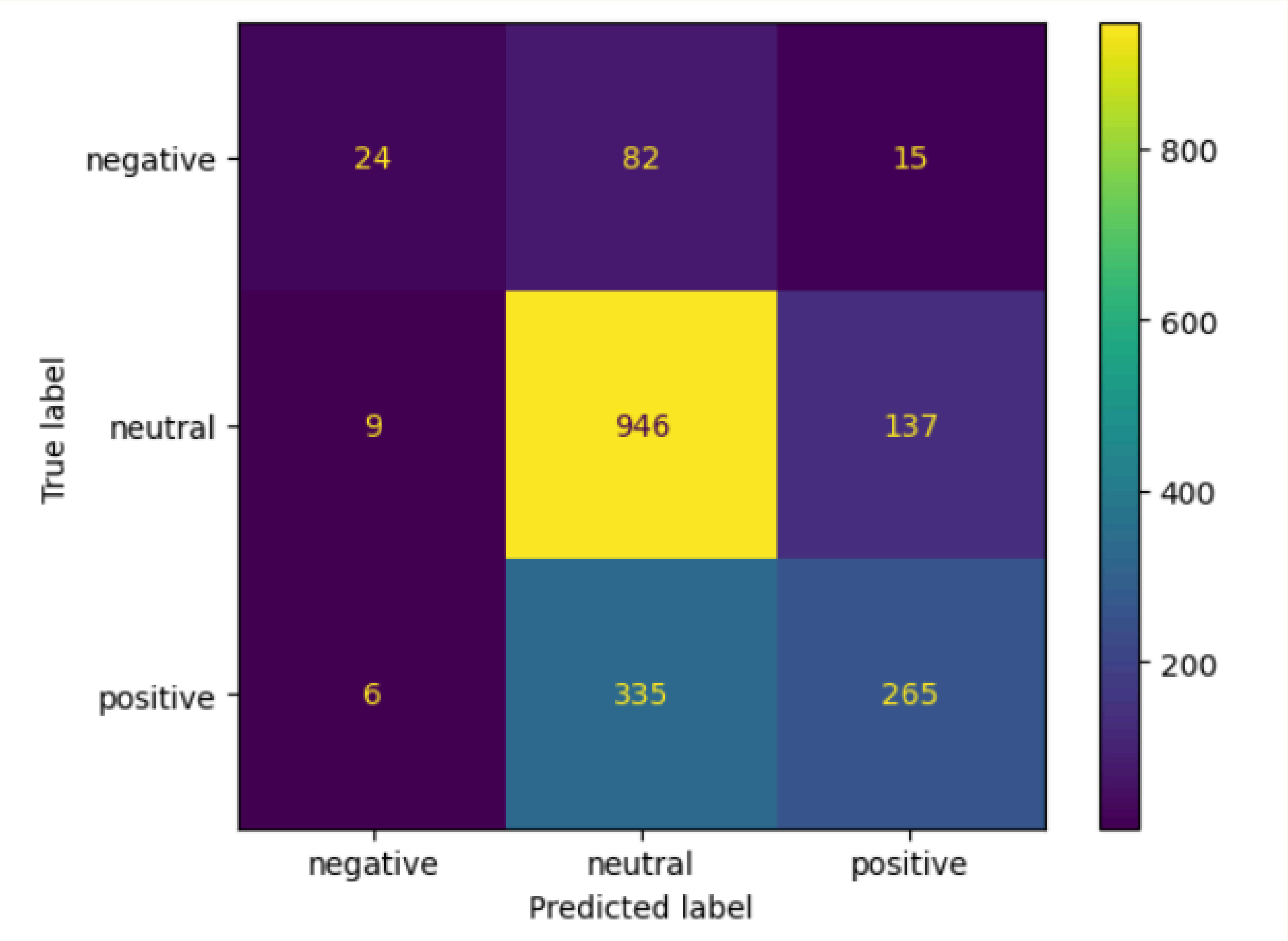
Opimizied Decision Tree Classifier



The confusion matrix and metrics indicate strong model performance. Accuracy (86.1%) shows overall correctness. Precision (84.4%) highlights the model's ability to identify true positives accurately, while recall (86.1%) shows sensitivity to positive cases. The F1 score (84.9%) balances precision and recall, indicating reliable classification.

The metrics and the inferred confusion matrix indicate that the Decision Tree Classifier with optimized parameters is achieving a good balance between accuracy, precision, and recall in classifying tweet sentiment. This suggests that the model is robust and can effectively distinguish between positive and negative sentiments.

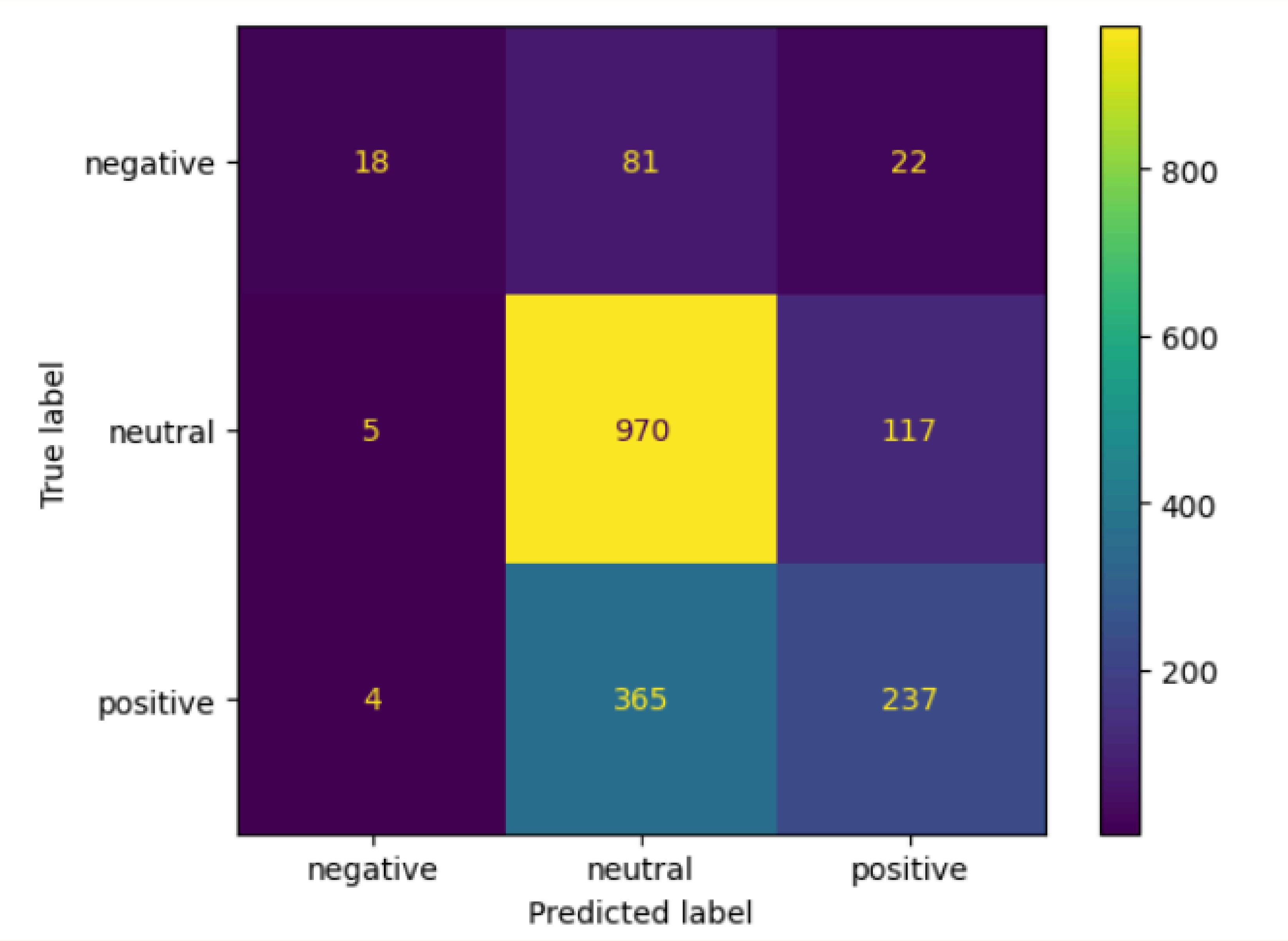
Random Forest Classifier



The confusion matrix and metrics reflect moderate performance for a random classifier. Accuracy (69.6%) shows decent overall correctness. Precision (69.2%) and recall (69.6%) are balanced, but lower than ideal. The F1 score (67.4%) suggests the model struggles with classifying certain labels effectively, especially positive and neutral categories.

The model's performance, as indicated by the accuracy, precision, recall, and F1-score, shows moderate results. While the model achieves a reasonable level of accuracy, there's room for improvement. The confusion matrix reveals that the model performs well on the "neutral" class but struggles more with the "positive" class, exhibiting a higher number of false negatives. To enhance performance, strategies such as hyperparameter tuning, feature engineering, and addressing potential class imbalances could be explored.

XGBoost



The XGBoost model achieves moderate accuracy (0.6701) and F1-score (0.6328), indicating a balance between precision and recall. The confusion matrix reveals that the model performs well on the "neutral" class but struggles more with the "positive" and "negative" classes, exhibiting a higher number of false positives and false negatives for these classes. To enhance performance, strategies such as hyperparameter tuning (e.g., adjusting learning rate, number of trees, maximum depth), feature engineering (e.g., creating new features from existing ones), and addressing class imbalance (e.g., oversampling, undersampling) could be explored.

CONCLUSION

The models, including Logistic Regression, Decision Tree, Random Forest and XGBoost, showed moderate performance with accuracies ranging from 67% to 88%. All models exhibited a good balance between precision and recall, indicating their ability to identify positive cases while minimizing false positives. The confusion matrices revealed that all models struggled to a varying degree with the "positive" and "negative" classes, exhibiting higher rates of false positives and false negatives for these classes compared to the "neutral" class.

RECOMMENDATIONS

1. Use the Logistic Regression model to monitor and respond to sentiment trends for improved customer engagement.

1. Address negative feedback to resolve consumer pain points and enhance products or services.

1. Benchmark against competitors by analyzing comparative sentiment data for Apple and Google.

1. Refine the model’s accuracy by improving feature engineering to handle nuanced tweets like sarcasm and abbreviations

@group7

THANK YOU

for your time and attention

Presented by Group Seven

Doreen Chepkong'a, Boniface Thuo, Linda Temoet
Grace Mwega, Micheal Munene