Twitter Sentiment Analysis on Apple and Google Products

Prepared by:

- Adrianna Ndubi
- Caleb Kipkech
- Gerald Muhia,
- Keith Kimani,
- Sarah Joshua

Overview and Business Understanding

- Apple: Premium hardware (iPhone, iPad, Mac) plus a cohesive software ecosystem.
- Google: Versatile products (Android OS, Pixel, Nest) plus a cross-platform integration.

Competition:

- Smartphones: iPhone vs. Pixel
- Tablets: iPad vs. Android
- Smart Home: HomeKit vs. Nest
- Ecosystems: Closed (Apple) vs. Open (Google)

Proposed Solution

Twitter Sentiment Analysis: Classify tweets (positive, negative, neutral) and analyze trends.

Key Stakeholders:

- Marketers: Campaigns based on sentiment.
- Product Teams: Customer-aligned development.
- Retailers: Sentiment-driven sales strategies.

Impact

- Enhanced brand loyalty.
- Improved product offerings.
- Targeted marketing strategies.

Problem Statement

- **Objective**:

Analyze public sentiment about Apple products using Twitter data.

Dataset:

Over 9,000 tweets rated by human annotators from CrowdFlower.

Goal:

Build an NLP model to classify tweets as positive, negative, or neutral.

Outcome

Provide actionable insights for Apple's marketing and product development.

Objectives

Main Objective:

- Build an NLP model to classify tweet sentiment.

Secondary Objectives:

- Compare sentiment for Apple vs. Google products.
- Identify key drivers of positive and negative sentiment.
- Monitor sentiment trends over time.
- Gather insights into customer preferences and opinions.

Metrics Of Success

Accuracy: ≥ 85%

Recall: ≥ 85%

Precision: ≥ 85%

F1 Score: ≥ 0.85 for each sentiment category.

Data Understanding

Source

CrowdFlower dataset (9,093 tweets from SXSW 2013).

Columns

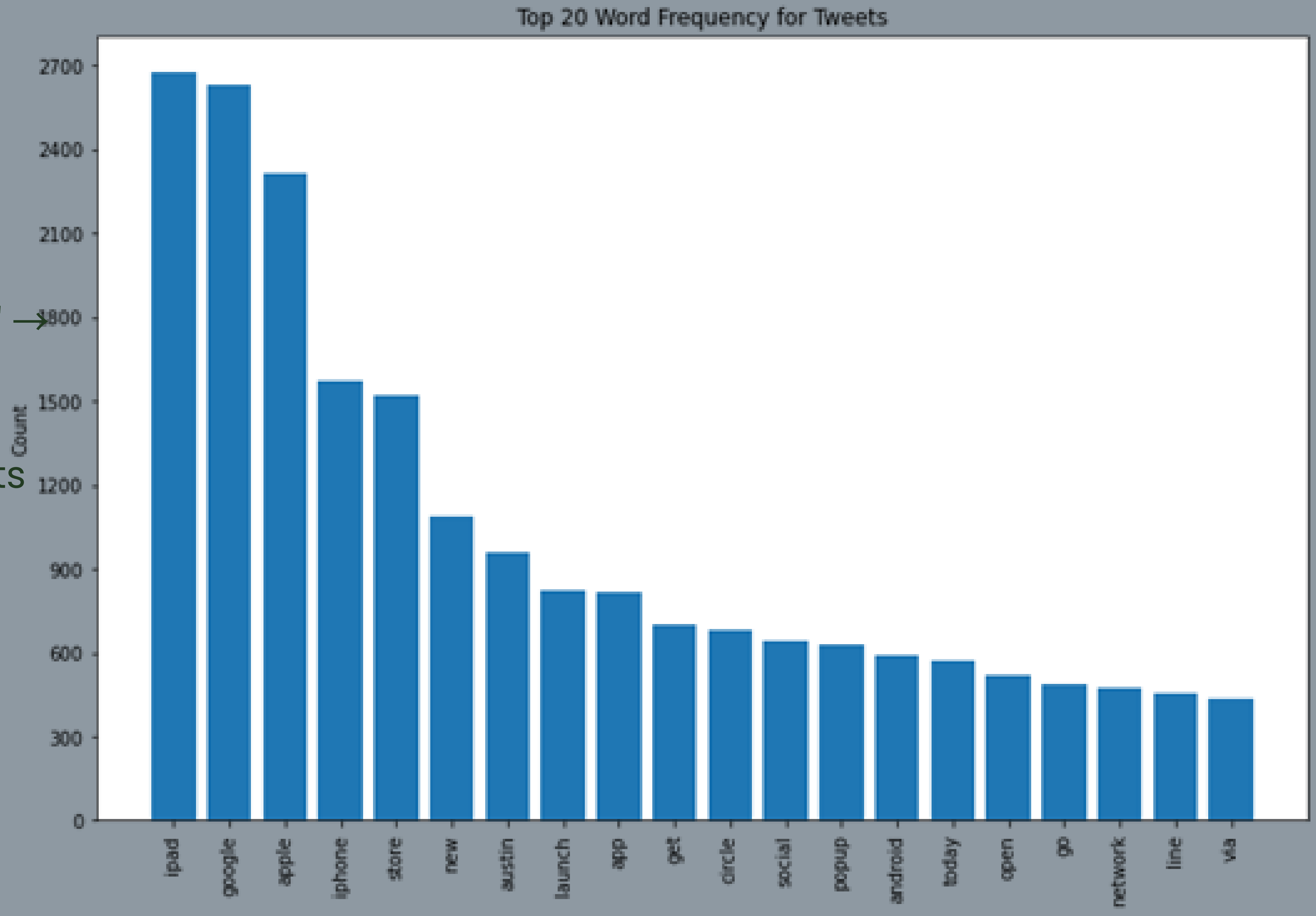
- tweet_text: Text of the tweet.
- Product_Name: Product being referred to.
- Sentiment_Type: Positive, negative, or neutral.

Data Preparation and Cleaning

The resulting top 20 words are as shown in the figure below:

Steps:

1. Renamed columns for clarity.
2. Handled missing values and duplicates.
3. Merged sentiment categories (e.g., "I can’t tell" → Neutral).
4. Text preprocessing:
   - Lowercasing, tokenization, hashtag/accents removal.
   - Punctuation and stop word removal.
   - Lemmatization and POS tagging.
5. Automated the preprocessing workflow.



Exploratory Data Analysis (EDA)

1. **Sentiment Distribution:**

- Neutral > Positive > Negative.

2.Brand Distribution:

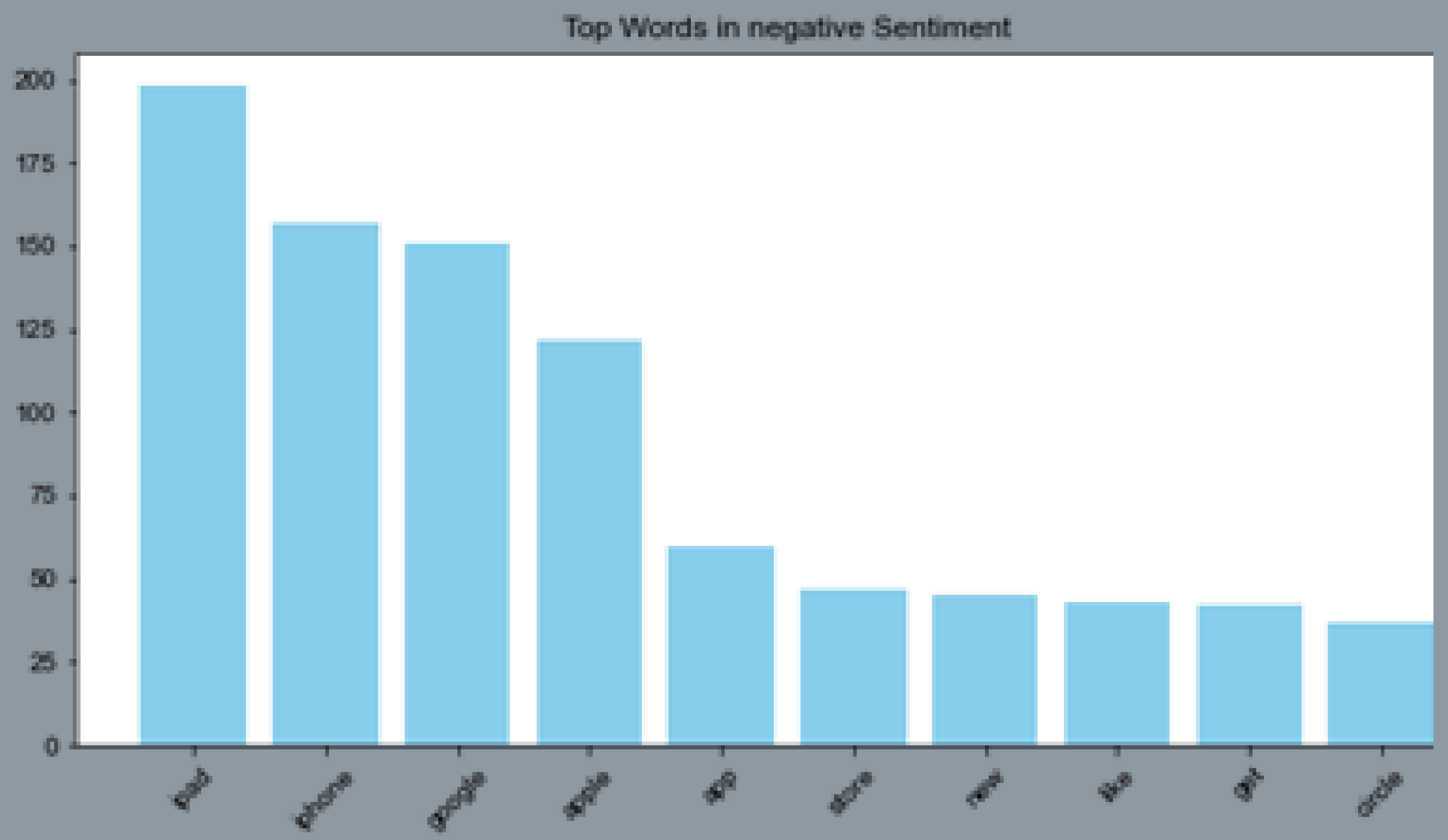
- Apple: 5,000+ tweets.
- Google: 3,000+ tweets.

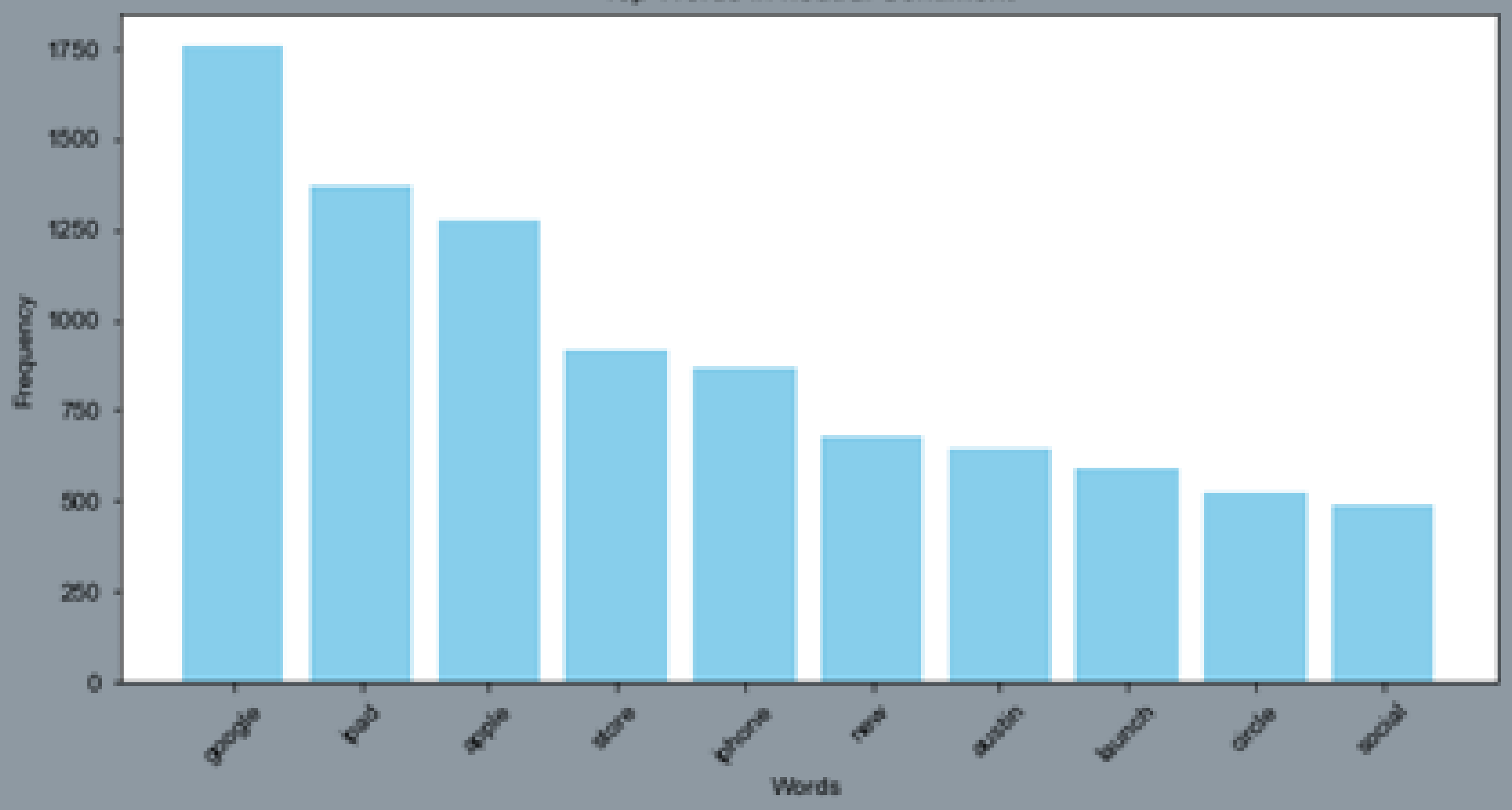
3. Sentiment by Brand:

- Apple has slightly more positive sentiment than Google.

4.Top Words by Sentiment:

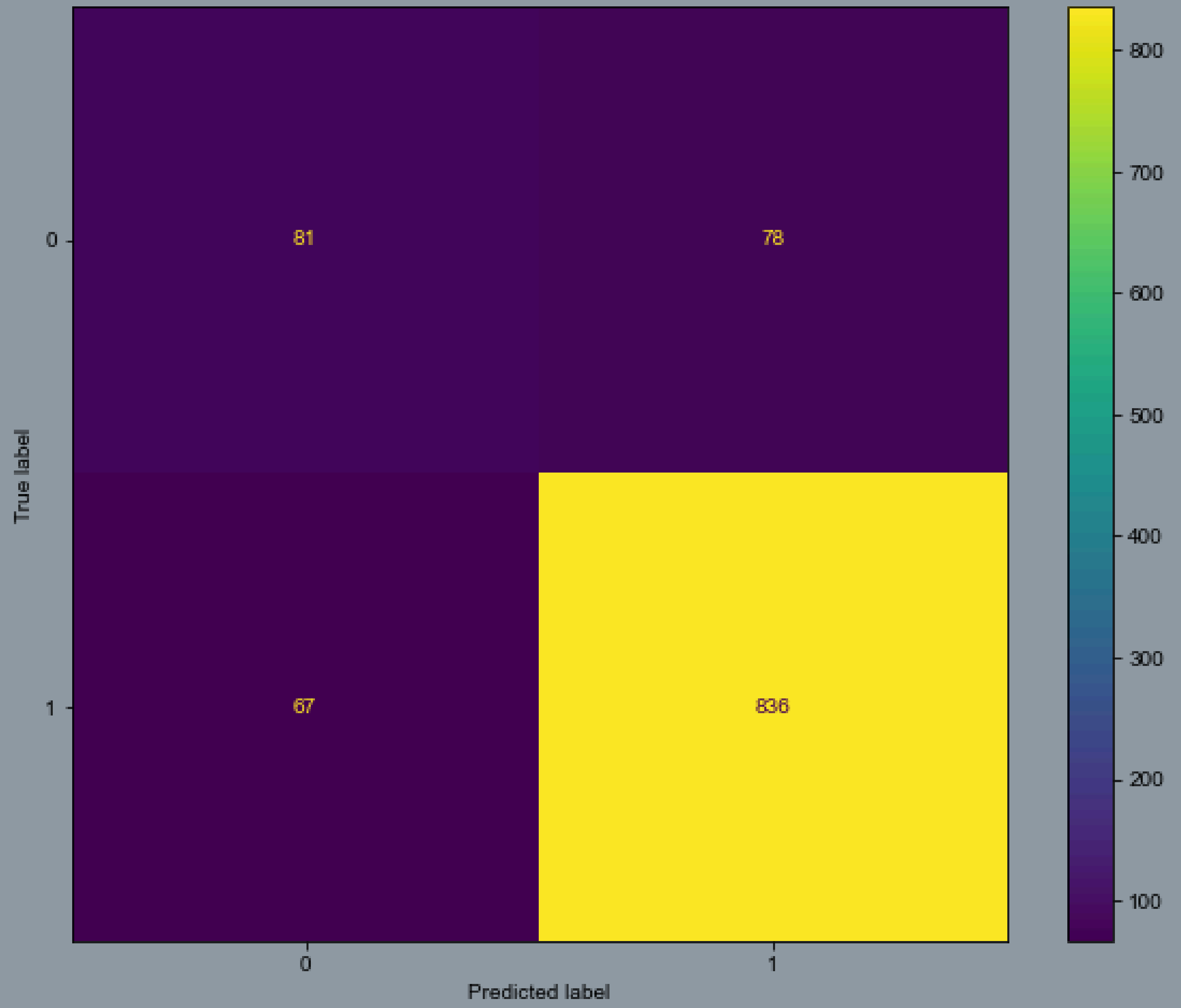
- Positive/Negative: "iPad" most frequent.
- Neutral: "Google" most frequent.





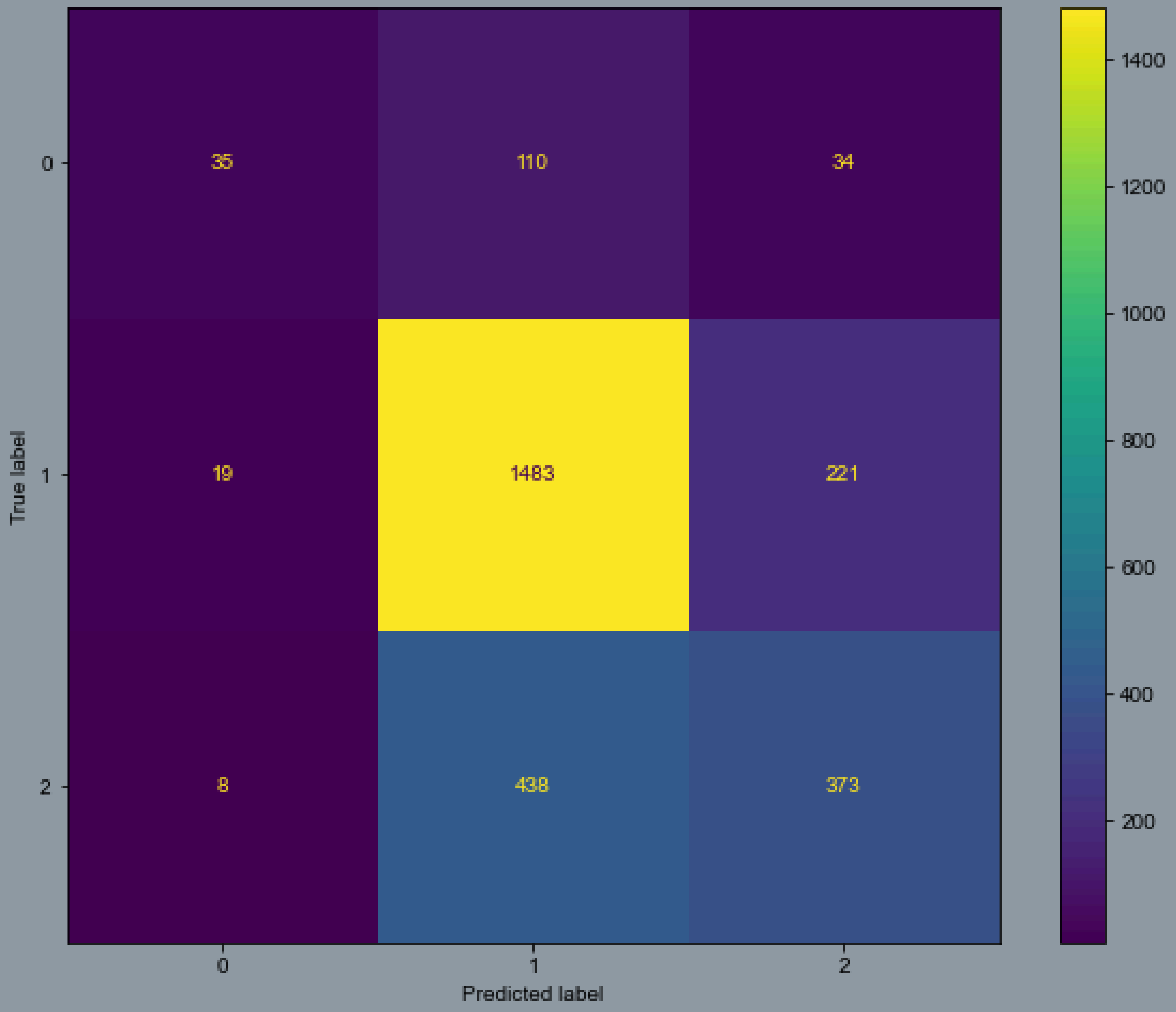
Modelling – Binary Classification

- Baseline Model: Multinomial Naïve Bayes (85% accuracy).
- Hyperparameter Tuning: Improved precision and recall to 88%.
- Oversampling (SMOTE): Decreased performance.
- Neural Network: Slightly lower performance than Naïve Bayes.
- Best Model: Tuned Multinomial Naïve Bayes without oversampling.



Slide 10: Modelling - Multi-class Classification

- Baseline Model: Multinomial Naïve Bayes (64% accuracy).
- Hyperparameter Tuning: Improved accuracy to 68%.
- XGBoost: Outperformed Naïve Bayes (70% accuracy).
- Oversampling (SMOTE): Slightly improved precision but decreased accuracy.
- Best Model: XGBoost without oversampling.



Model Evaluation

Binary Classification

- Oversampling worsened performance.
- Model tuning improved Multinomial NB, not Neural Networks.
- Multinomial NB outperformed Neural Networks.
- Chosen model: Multinomial NB without oversampling.

Multi-class Classification:

- Oversampling worsened performance in addressing class imbalance.
- Multiclass classification metrics were lower than binary classification.
- Model tuning improved results across all models.
- XGBoost outperformed Multinomial NB for multiclass sentiment analysis.
- Final choice: XGBoost without oversampling, but metrics fell short of success criteria.

Key Observations:

- Class imbalance significantly impacted performance.
- Oversampling did not consistently improve results.
- Minority classes (negative sentiment) were harder to classify.

Conclusion

Key Findings:

- Neutral sentiment dominates across both brands.
- Apple has a higher proportion of positive sentiment.
- Challenges in classifying minority sentiment classes (negative).

Best Models:

- Binary Classification: Tuned Multinomial Naïve Bayes.
- Multi-clas Classification: XGBoost.

Next Steps:

- Explore advanced techniques (e.g., deep learning).
- Address class imbalance more effectively.
  Incorporate external data sources for richer insights.

Thank You!