Natural Language Processing Project

Sentiment Analysis of
Tweets Using NLP

Business Understanding

Overview

- Role of social media in capturing user sentiments.
- Use of machine learning and NLP to categorize sentiments.
- Objective: Analyze sentiments towards Apple and Google products from Twitter data.

Business Understanding

Objective:

- Collect Twitter data related to Apple and Google product reviews.

- Preprocess and perform sentiment analysis.

- Derive insights to optimize product assortment and align with customer preferences.

Business Understanding

Challenges:

- Lack of systematic methods for sentiment analysis on platforms like Twitter.

- Need for data-driven stocking decisions.

Project aims:

Enhance understanding of customer opinions to improve stocking decisions and customer satisfaction.

Business Understanding

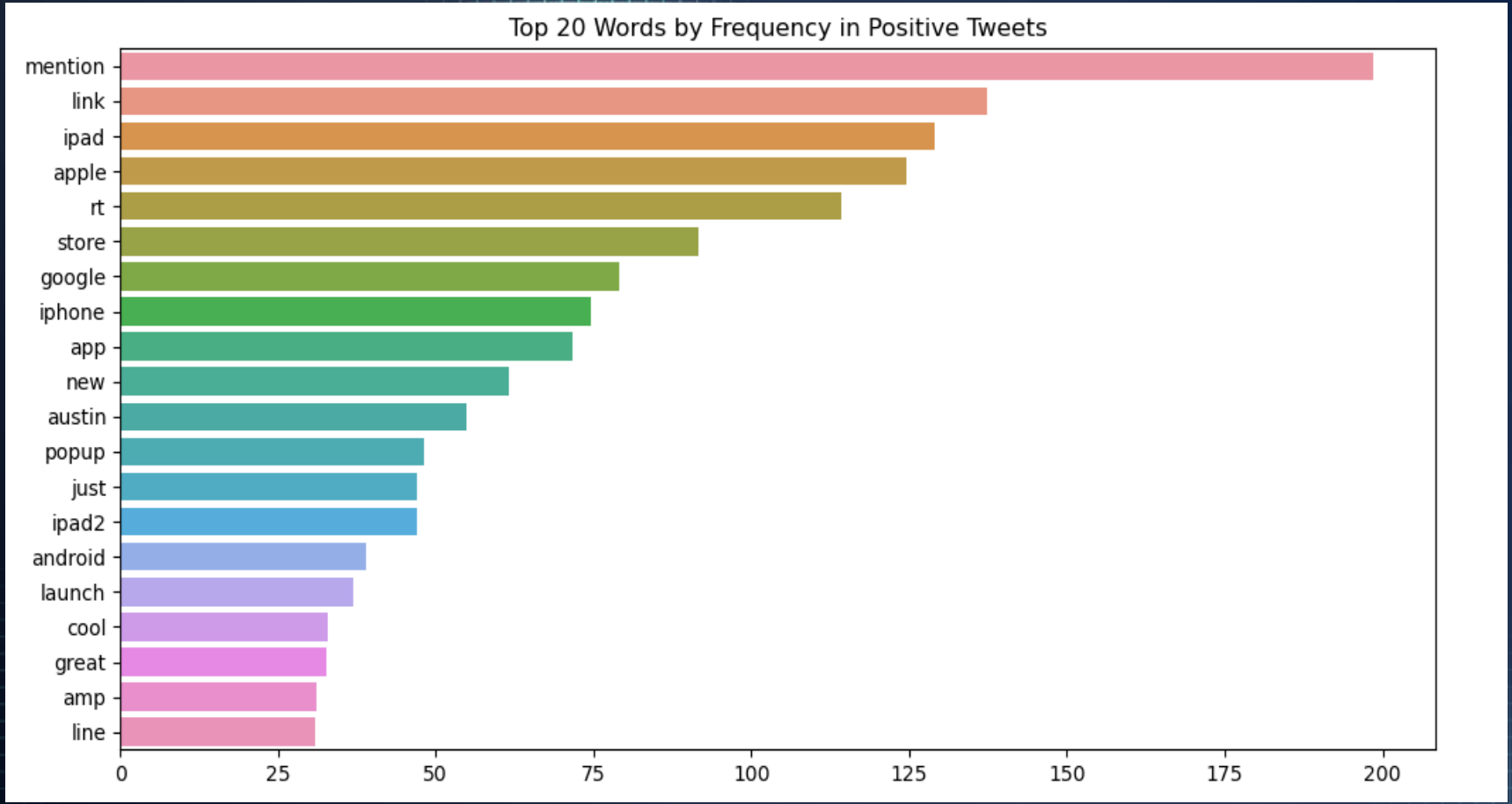
Business Problem:

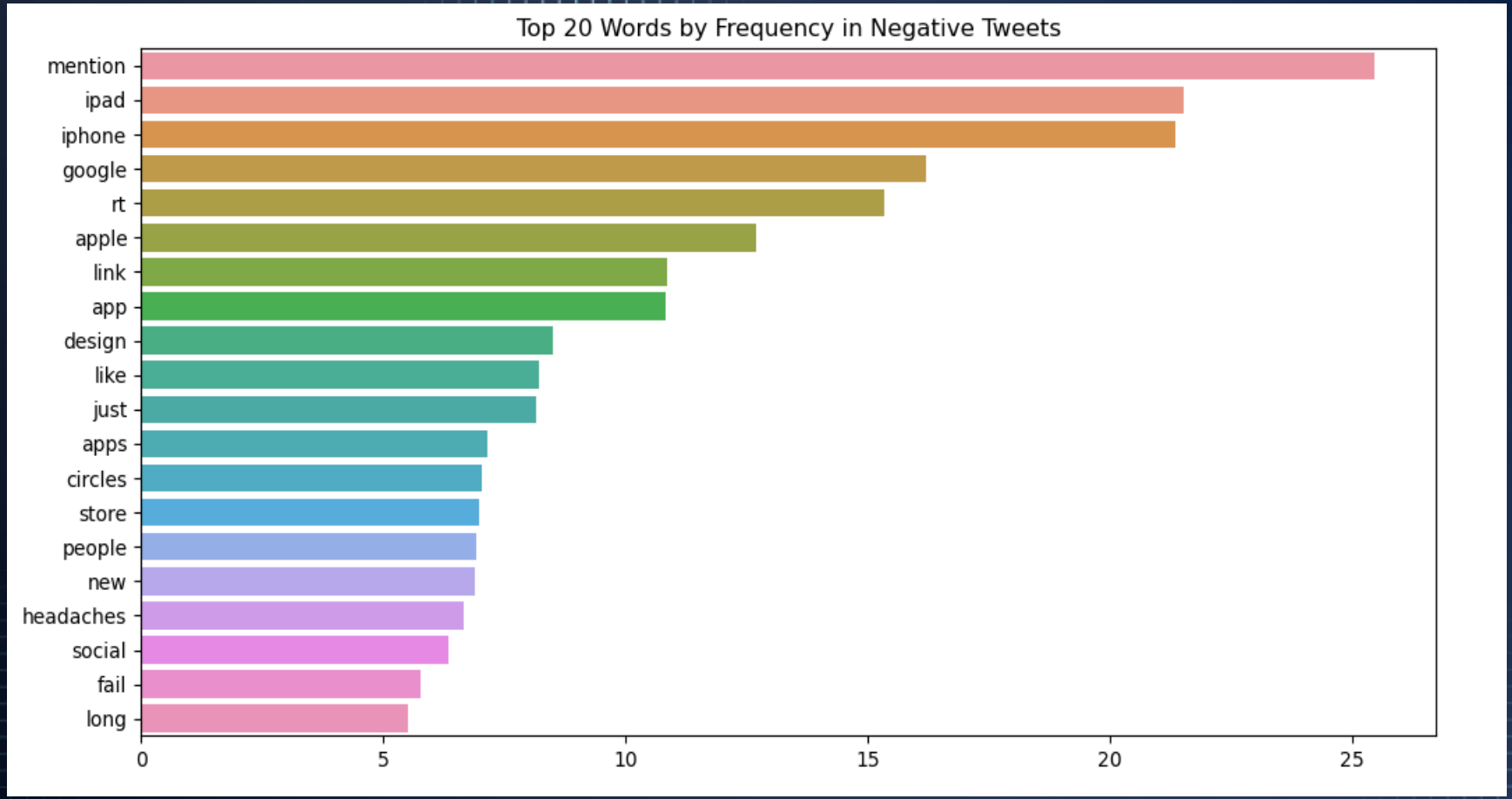
- Ensure a positive customer experience by accurately predicting sentiment.

- Importance of correctly classifying sentiments to avoid misinterpreting customer feedback.

Data Preprocessing

- **Data Cleaning:**
  - Removing Duplicates: Ensuring each tweet is unique.
  - Handling Missing Values: Strategies for dealing with incomplete data.
- **Text Preprocessing:**
  - Tokenization: Breaking down text into individual words or tokens.
  - Stop-word Removal: Eliminating common words that do not contribute to sentiment (e.g., "and," "the").
  - Lemmatization/Stemming: Reducing words to their base or root form.
- **Data Transformation:**
  - Converting text to lowercase.
  - Removing special characters and URLs.


Top 20 Words by Frequency in Positive Tweets
mention
link
ipad
apple
rt
store
google
iphone
app
new
austin
popup
just
ipad2
android
launch
cool
great
amp
line
0
25
50
75
100
125
150
175
200


Top 20 Words by Frequency in Negative Tweets
mention
ipad
iphone
google
rt
apple
link
app
design
like
just
apps
circles
store
people
new
headaches
social
fail
long
0
5
10
15
20
25

Model Training and Evaluation

Model Training:

- Splitting Data: Dividing data into training and test sets.
- Feature Extraction: Converting text into numerical features using techniques like TF-IDF or word embeddings.
- Training Process: Details on model training, including parameters and iterations.

Cont ...

Evaluation Metrics:

- Accuracy: Proportion of correctly classified tweets.
- Precision, Recall, F1-Score: Metrics to evaluate the performance of the model in detail.
- Confusion Matrix: Visual representation of the model's performance.

Cross-Validation:

- K-Fold Cross-Validation: Ensuring the model's robustness by training and validating on different data splits.

Insights and Recommendations

Actionable Insights:

- Product Improvement: Areas where customers expressed dissatisfaction, suggesting potential improvements.
- Marketing Strategy: Insights into how customers perceive marketing campaigns.
- Customer Support: Identifying common customer concerns that can be addressed through better support.

Recommendations:

- Tailoring advertising strategies based on sentiment analysis.
- Enhancing product features that received positive feedback.
- Addressing common issues to improve customer satisfaction.

The End

Thank You!