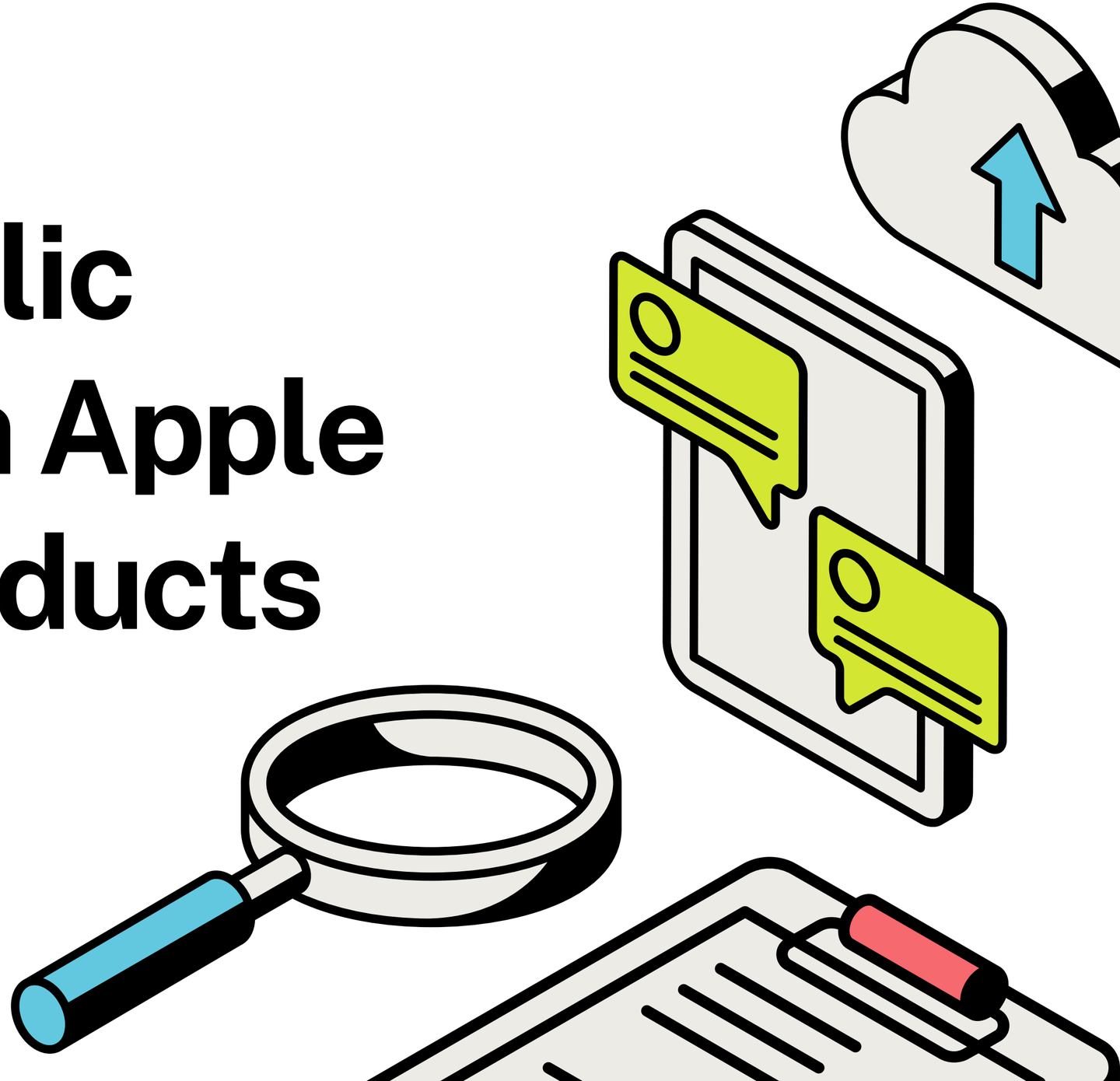
GROUP 4- DSF_PT11

Tracking Public Sentiment on Apple & Google Products with NLP

Mohamed Abdi Sheikh

George Kariuki

Deborah M Omondi

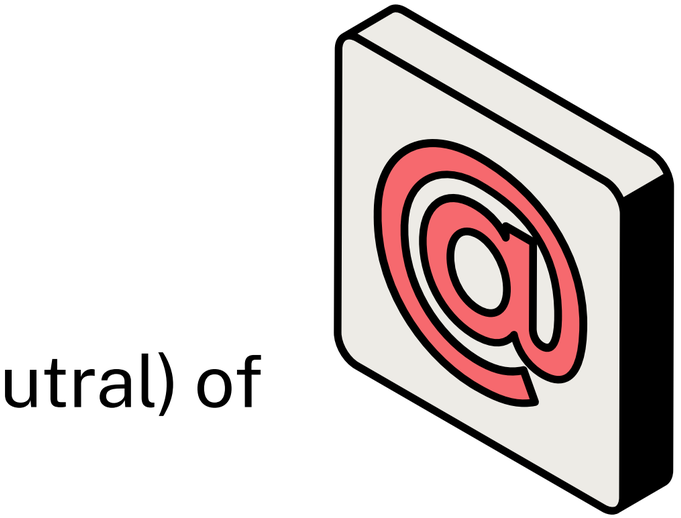


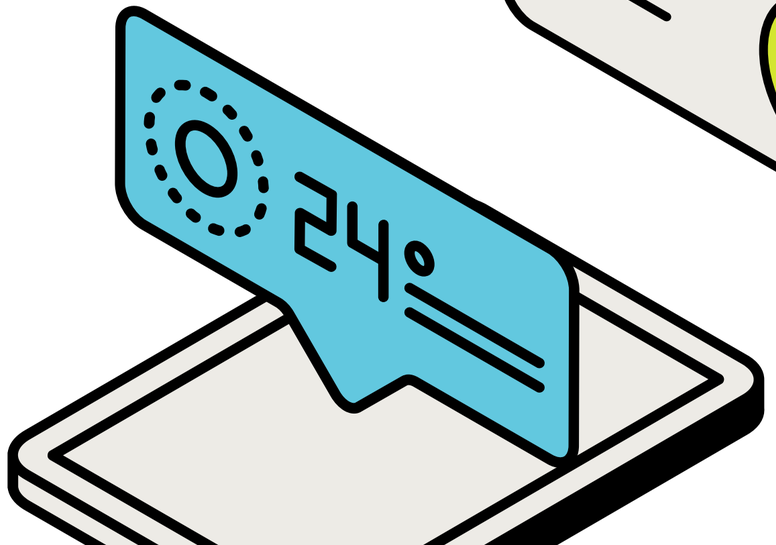
Problem Definition

Goal: Predict sentiment (positive, negative, neutral) of Tweets about Apple/Google products.

Why NLP? → Text data dominates customer feedback.

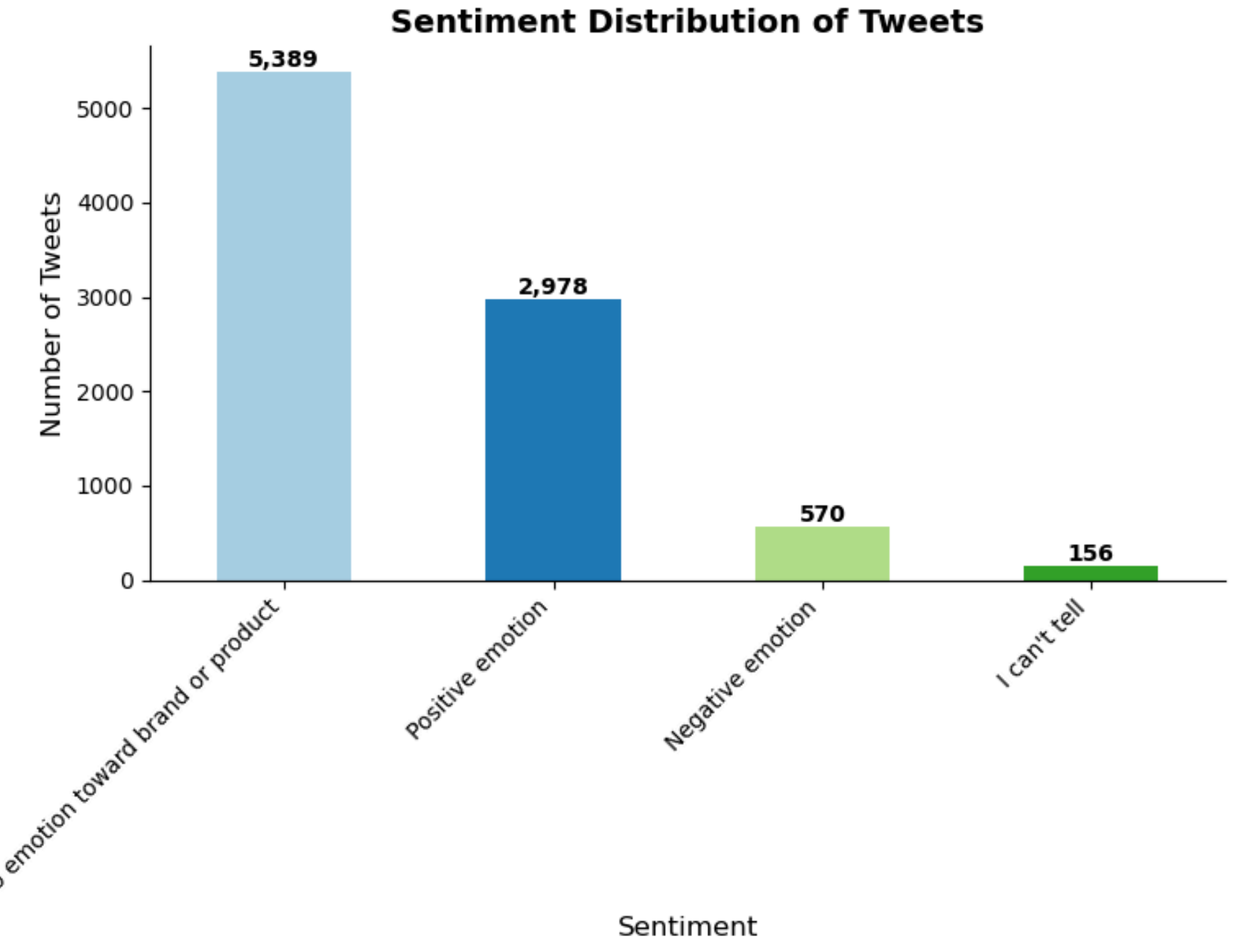
Why is it important? → Brand monitoring, customer insights, marketing optimization.





Dataset Overview

- **Source:** CrowdFlower via Data.World.
- **Size:** ~9,000 labeled Tweets.
- **Sentiment classes:**
  -No emotion
  - Positive
  -Negative
  - Neutral
- **Challenges:** class imbalance, noisy/short text, emojis/slang.



Data Preprocessing

A

Cleaning steps:

- **Lowercasing**
- **Removing punctuation, URLs, mentions, stopwords**
- **Tokenization & stemming/lemmatization**

B

Feature Engineering:

- **Bag-of-Words**
- **TF-IDF representations**

C

Handling imbalance: SMOTE (Synthetic Minority Oversampling).

Modeling Approach

Baseline models:

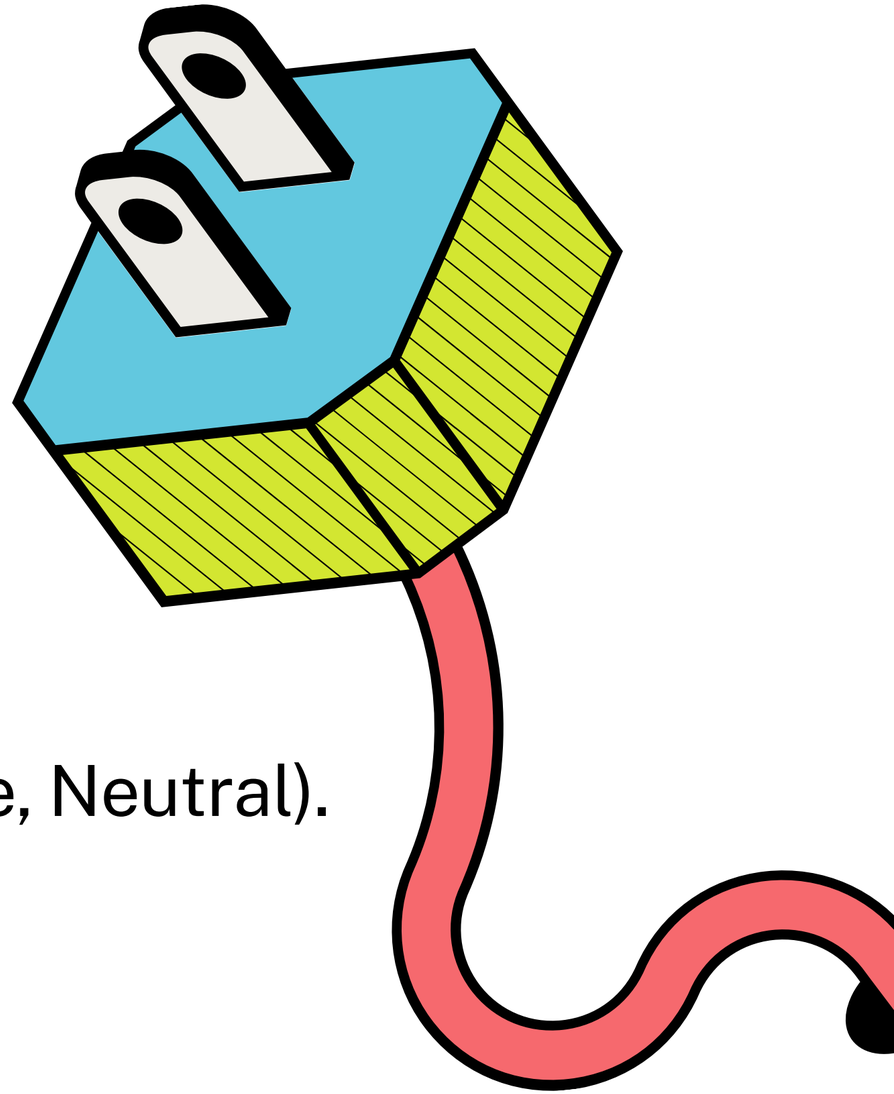
- Naïve Bayes
- Logistic Regression

Additional models:

- Support Vector Machines (SVM)
- Ensemble methods (Voting Classifier)

Training setups:

- Binary classification (Positive vs. Negative)
- Multiclass classification (Positive, Negative, Neutral).

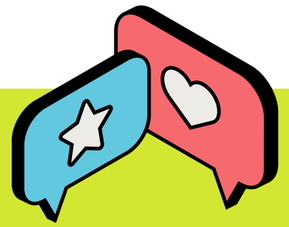


Evaluation Metrics

- **Binary classification:** Accuracy, Precision, Recall, F1-score.
- **Multiclass classification:** Weighted average F1, Confusion Matrix.
- **Why F1?** → Balances precision & recall, important for imbalanced data.

Results: Binary Classification

Logistic Regression



88% Accuracy

Naive Bayes

82% Accuracy

SVM

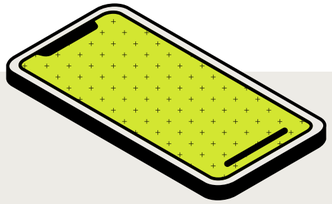
85% Accuracy

Results: Multiclass Classification

Ensemble model

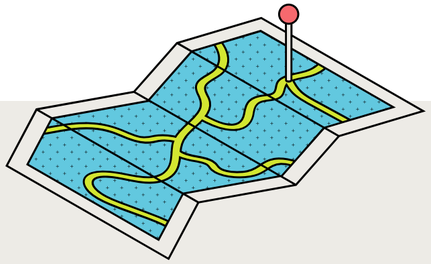
68% accuracy

Confusion Matrix



Neutral class hardest to predict

Insights



Negative Tweets best detected, neutral tweets often confused with positive/negative.

Recommendations

1	Deploy dashboard for monitoring
2	Prioritize negative sentiment alerts
3	Retrain model regularly
4	Explore advanced AI (BERT)
5	Add explainability (highlight keywords)

Thank you.