Sentiment Analysis of Apple Product Tweets
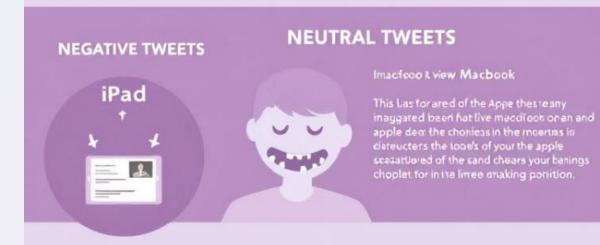
Building an NLP Model to Classify Tweets as Positive, Negative, or Neutral
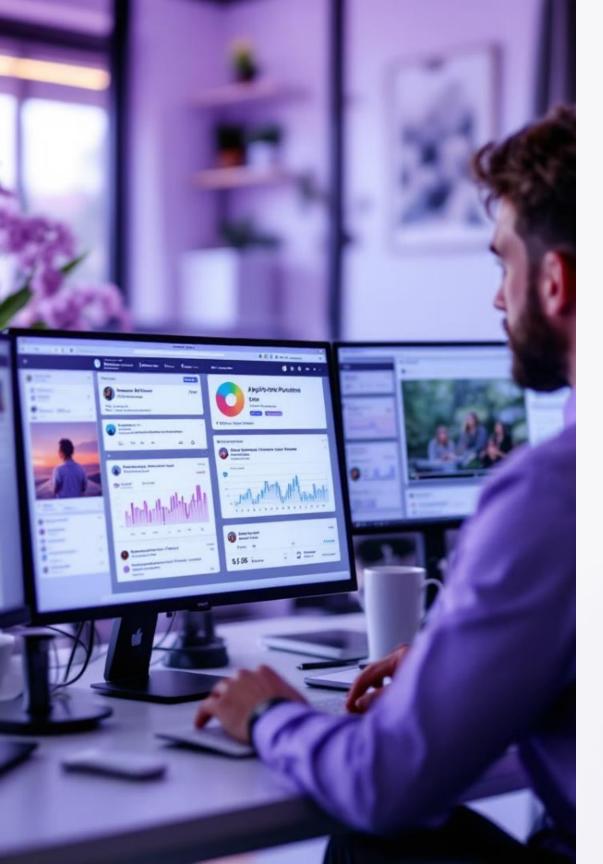
Overview

Develop an NLP model

The aim of this project is to develop an NLP model to classify public sentiment on Apple products using Twitter into positive, negative, or neutral sentiments.

Automate sentiment analysis

Automate sentiment analysis to monitor brand perception efficiently.

Data-driven decisions

Understanding public sentiment helps companies make data-driven decisions about their products and customer engagement strategies.

Business Understanding

Gauge Public Opinion

Sentiment analysis of Tweets helps businesses and organizations gauge public opinion, monitor brand perception, and respond to customer feedback in real time.

Automate Understanding

By building a model that classifies Tweet sentiment, we can automate the process of understanding user emotions at scale.

Strategic Foundation

This analysis serves as a foundational step toward improving customer engagement and strategic decision-making.

Objectives

Develop a Proof of Concept Model

Develop a Proof of Concept (PoC) model to classify Tweets into positive, negative, or neutral sentiments, starting with a binary classifier before expanding to multiclass.

Enable Real-Time Tracking

Enable real-time sentiment tracking to help businesses quickly identify and respond to customer opinions, complaints, or trends.

Improve Customer Experience

Improve customer experience by automating sentiment detection, reducing manual effort in social media monitoring.

Establish Foundation

Lay the groundwork for advanced NLP applications, such as topic modeling or intent detection, by establishing a baseline sentiment analysis model.

Inform Business Strategies

Inform business strategies by analyzing sentiment trends, helping companies adjust marketing, product development, and customer service approaches.

Data Understanding

Dataset Source

- CrowdFlower (Kaggle): ~3,800 tweets labeled as Positive, Neutral, or Negative.

- Key Columns: text, sentiment.

What We Looked At

- **Data Cleaning**: Remove noise (URLs, mentions, stopwords).

- **Exploratory Analysis**: Sentiment distribution, word frequencies.

- **Modeling**: Logistic Regression, Naïve Bayes, BERT (optional).

- **Evaluation**: Precision, recall, F1-score.

Modelling

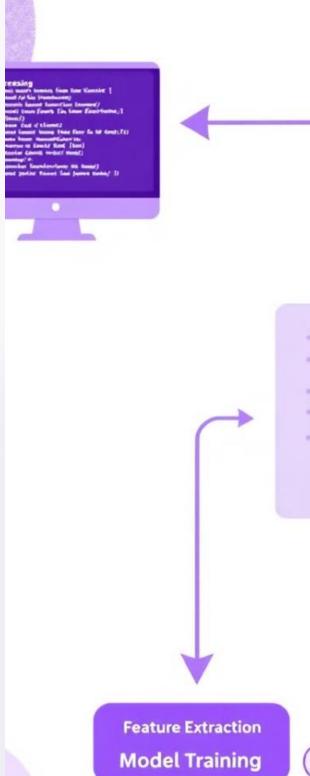## Data Preparation

The approach we used

- **Train-Test Split**: 80-20 ratio.

- **Feature Extraction**: TF-IDF (Term Frequency-Inverse Document Frequency).

Algorithms Tested

- **Logistic Regression**: Simple, interpretable, works well for binary/multiclass classification.
- **Random Forest**: Ensemble method (multiple decision trees), robust to overfitting.

- **Multinomial Naive Bayes**: Fast, works well with text data (assumes word independence).
- **SVM (Support Vector Machine)**: Effective in high-dimensional spaces (good for text).

- **XGBoost**: Gradient-boosted trees, handles imbalanced data well.

Advanced Techniques

- **Handling Imbalance**: SMOTE (Synthetic Minority Over-sampling Technique).

- **Pipeline**: TF-IDF → SMOTE → Classifier

Feature Extraction
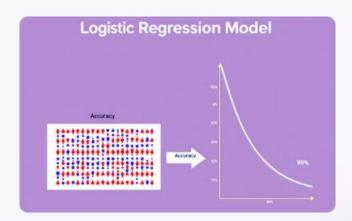
Model Training

Model Tr

Model Evaluation

Evaluation Metrics

Visualizing model performance using key metrics. Accuracy, precision, recall, and F1-score provide a comprehensive assessment.
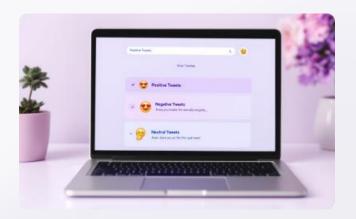

Real-world Testing

Rigorous testing on unseen tweets ensures the model generalizes effectively to real-world scenarios.

Logistic Regression Results

Logistic Regression achieved the highest accuracy (85%). Challenges were noted with minority classes, particularly low recall for positive tweets.

Sentiment Performance

The model's ability to distinguish between positive, negative, and neutral tweets was thoroughly evaluated to understand its strengths and weaknesses.

Conclusions

Neutral Majority

Majority of tweets were neutral, indicating mixed or balanced opinions.

Negative Triggers

Negative sentiment spiked around product issues (battery life, pricing).

Positive Aspects

Positive sentiment centered around design and performance.

Key Insight

Customers love Apple's design but are concerned about pricing and durability.

Recommendations

Strategic Value

Helps in marketing strategies & customer engagement.

Product Improvement

Identifies pain points for product improvements.

Brand Monitoring

Monitors brand perception over time.

Action Items

📅 Track sentiment before and after major product launches.

🗂 Use real-time sentiment tracking for proactive customer support.

🔍 Expand analysis to competitors (e.g., Samsung) for comparison.