Natural Language Processing (NLP)Machine Learning Classification Model to Predict Sentiments of Apple Products on Twitter

Recommendations to Apple on Best Model to Identify Negative Sentiments on Twitter

Contents

- Overview
- Business and Data Understanding
- Modelling Process
- Evaluation Process
- Recommendations
- Next Steps
- Thank you !

Overview

- This project examines a dataset from https://data.world/crowdflower/brands-and-product-emotions that classifies tweets about Apple and Google products as positive, negative or neutral.
- By analyzing these sentiments from the tweets about their products and that of their competitor, Apple can tap into a wealth of authentic feedback that traditional surveys or feedback forms might miss.
- This immediate access to customer sentiment will allow them to swiftly identify trends, preferences, and potential issues, allowing for proactive engagement and timely adjustments to strategies.

Business Objectives

The main objectives of this project was to train various baseline and ensemble models using the provided dataset to accurately predict the sentiment on a tweet, and to analyze how Apple products are rated by customers compared to those of their competitor Google.

This was done by doing the following:

- Analyzed the distribution of negative and positive tweets by company; this was crucial in assessing how the sentiments of Apple products compare to those of Google products (competition landscape analysis)
- Trained, tuned and evaluated several classification models to identify positive, negative and neutral sentiments on previously unseen tweets.
- Trained, tuned and evaluated several classification models to identify **negative** sentiments on previously unseen tweets.

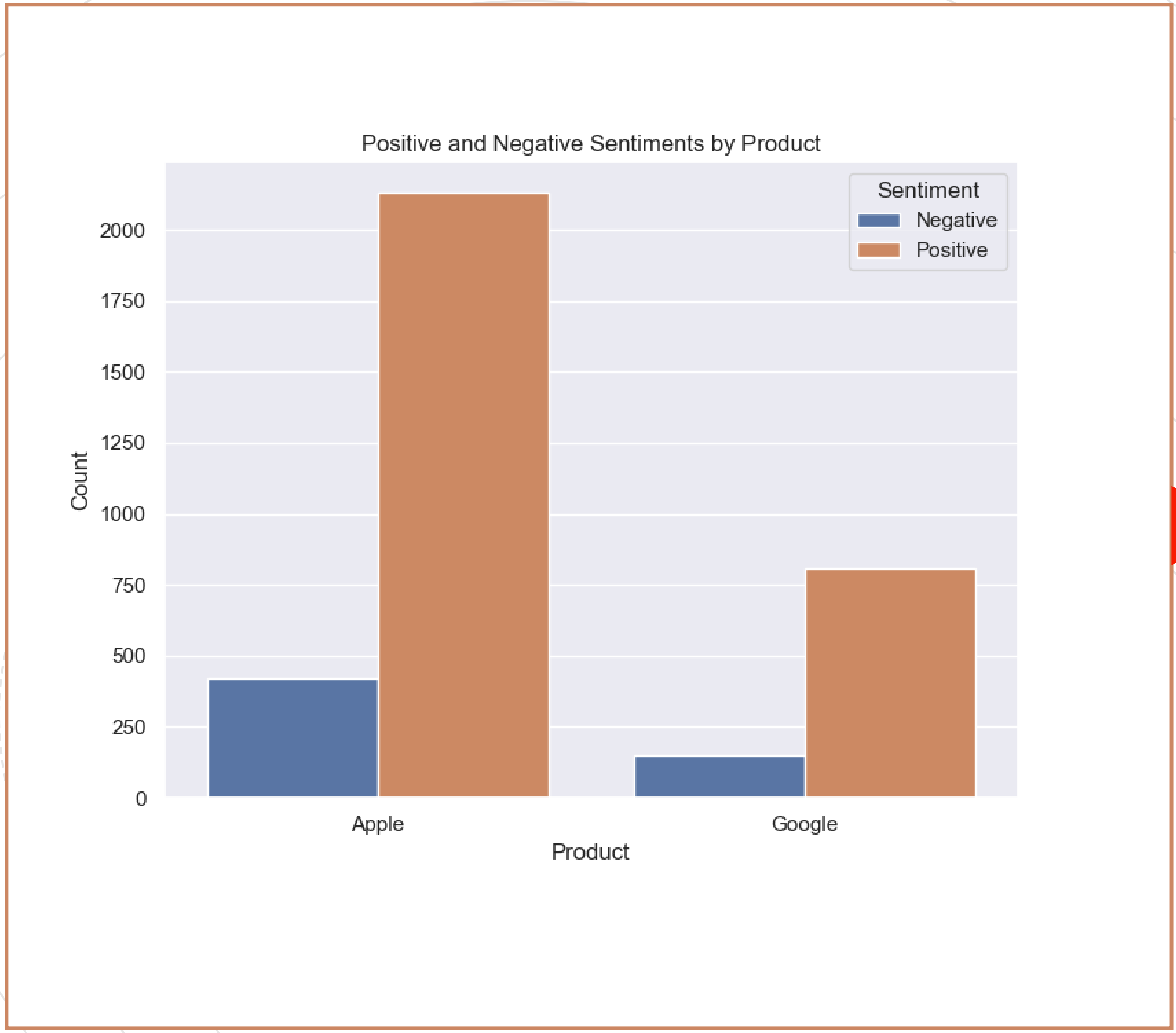
From the above iterative modelling process, provide to Apple the most optimal model to deploy on future (new) tweets to identify negative sentiments on their products.

The Modeling Process

- Machine learning algorithms, best suited for an NLP classification task were trained on the dataset provided to see which works best. The models trained were:
  - Logistic Regression: A modelling method that predicts the likelihood of outcomes
  - Naïve Bayes: A simple but powerful technique based on probability
  - Ensemble Models:: These combine multiple models to improve performance and include:
    - Random Forest: A bunch of decision trees working together.
    - AdaBoost: Boosting weaker models to make them stronger.
    - XGBoost: A highly efficient and powerful boosting method.
- An iterative process was used on the split dataset with distinct training and test sets. Each of the model was trained on the training set and tested on the test set.
- Starting with a baseline model logistic model, as a starting point to compare other models against, the models are systematically optimized with machine learning techniques suitable for each model.

The Evaluation Process

- In every step of the iterative modelling process, the following evaluation metrics are used to gauge the performance of each model:-
- Classification_report: Provides a summary of important performance metrics like:
  - Accuracy: How often the model gets the right answer.
  - Precision: How many of the predicted positive results are actually positive.
  - Recall: How many of the actual positive results are correctly predicted by the model.
  - F1-score: A balance between precision and recall.
- Confusion_matrix: It provides a visual summary of the prediction results by showing the count of true positives, true negatives, false positives, and false negatives.
- The model with the best metrics at the end of the iterative process will be recommended for deployment.



Objective # 1

Sentiments Analysis and Competition Landscape Evaluation

- Popularity and Sentiment Balance: Apple products are more popular but also have higher negative sentiments.
- Apple should monitor and address negative sentiments by enhancing customer service, improving product quality, and engaging with users on social media.to maintain its market position.

Model Evaluation

Objective # 2

Top 3 Best Performing Classification Models to Identify Positive, Neutral, and Negative Classes

- Baseline Logistic Regression Model -accuracy of 65%, while the tuned Logistic Regression and Random Forest Model had an accuracy of 66% and 67%, respectively. Comparing the metrics for the negative class:
- Precision: Random Forest is better at avoiding false positives for negative tweets.
- Recall: Baseline Logistic Regression captures a higher percentage of actual negative tweets.
- F1-Score: Baseline Logistic Regression offers a balanced approach with better recall.
- **The Baseline Logistic Model is the better model for identifying the three classes**
- Sub-optimal performance can be attributed to class imbalance. Although SMOTE was used to oversample the minority class, the synthetic data did not significantly enhance model performance.

Classification Report - Baseline Logistic Regression

	precision	recall	f1-score	support
Negative	0.37	0.48	0.41	189.00
Neutral	0.74	0.72	0.73	1612.00
Positive	0.56	0.56	0.56	880.00
accuracy	0.65	0.65	0.65	0.65
macro avg	0.56	0.58	0.57	2681.00
weighted avg	0.66	0.65	0.65	2681.00

Classification Report - Tuned Logistic Regression

	precision	recall	f1-score	support
Negative	0.42	0.39	0.40	189.00
Neutral	0.74	0.74	0.74	1612.00
Positive	0.57	0.58	0.57	880.00
accuracy	0.66	0.66	0.66	0.66
macro avg	0.57	0.57	0.57	2681.00
weighted avg	0.66	0.66	0.66	2681.00

Classification Report - Random Forest

	precision	recall	f1-score	support
Negative	0.66	0.21	0.32	189.00
Neutral	0.70	0.84	0.76	1612.00
Positive	0.60	0.48	0.53	880.00
accuracy	0.67	0.67	0.67	0.67
macro avg	0.65	0.51	0.54	2681.00
weighted avg	0.67	0.67	0.66	2681.00

Model Evaluation

Objective # 3

Evaluation of the Top 3 Best Performing Classification Models to Identify the Negative Class

- In our quest to develop a model with a higher recall for the negative class, we undertook the following steps:
  - Class Consolidation: We combined the Neutral and Positive classes into a new class labelled 'Other'.
  - Resampling: We built a model with a resampled reduced subset of the new class to address the class imbalance between Other and Negative classes
  - Model Training: We trained both baseline and tuned Logistic Regression models, along with three Ensemble models.
  - **Baseline Logistic Regression model provides the highest recall for the negative class at 75%, which is crucial for identifying negative sentiments accurately.**

Classification Report - Baseline Logistic Regression

	precision	recall	f1-score	support
Negative	0.69	0.75	0.72	183.00
Other	0.78	0.73	0.76	228.00
accuracy	0.74	0.74	0.74	0.74
macro avg	0.74	0.74	0.74	411.00
weighted avg	0.74	0.74	0.74	411.00

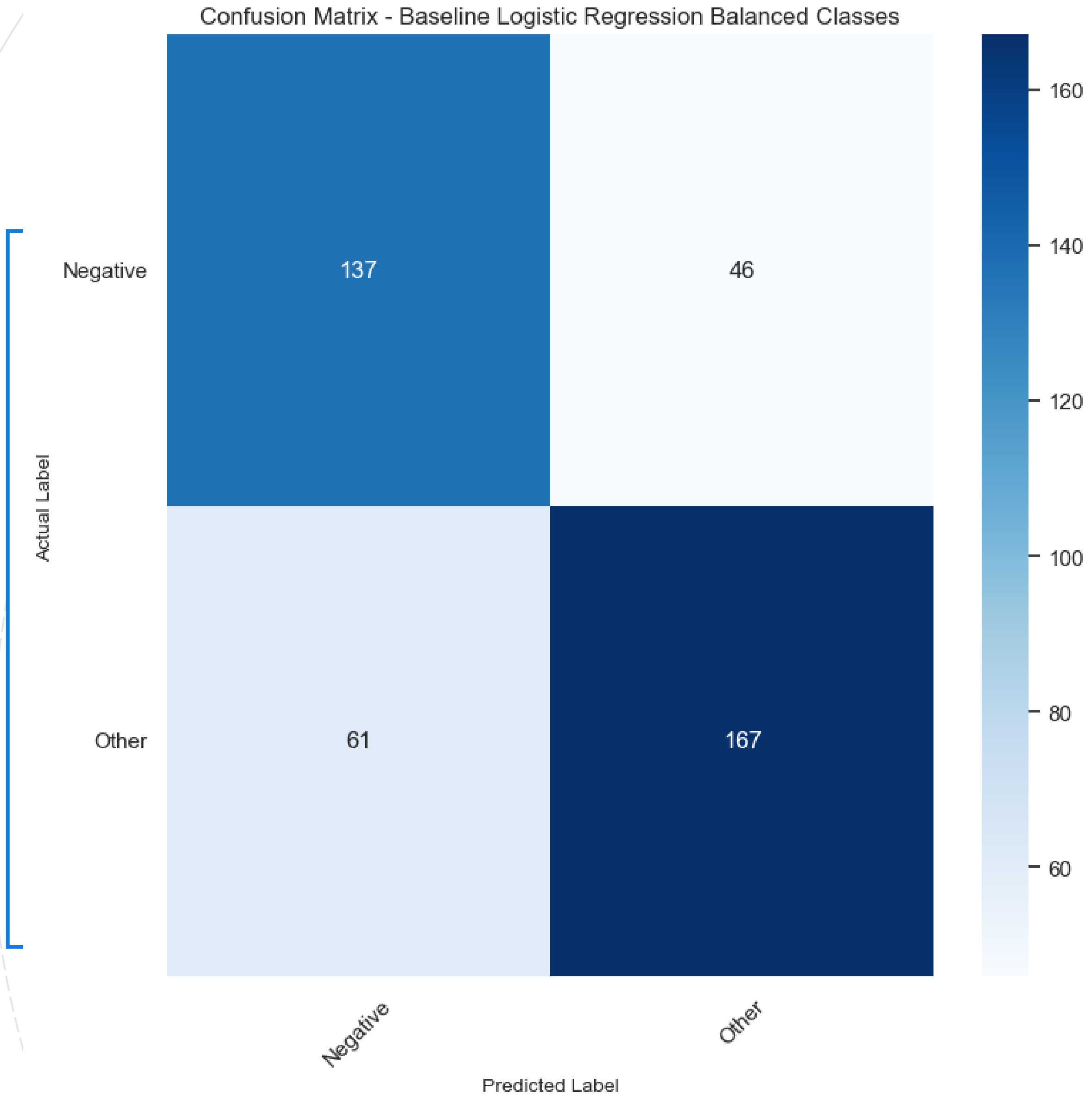
Classification Report - Logistic Regression

	precision	recall	f1-score	support
Negative	0.71	0.73	0.72	183.00
Other	0.78	0.76	0.77	228.00
accuracy	0.74	0.74	0.74	0.74
macro avg	0.74	0.74	0.74	411.00
weighted avg	0.75	0.74	0.74	411.00

Classification Report - Random Forest

	precision	recall	f1-score	support
Negative	0.76	0.58	0.66	183.00
Other	0.72	0.86	0.78	228.00
accuracy	0.73	0.73	0.73	0.73
macro avg	0.74	0.72	0.72	411.00
weighted avg	0.74	0.73	0.73	411.00

Recommendations:



The focus of the business should be in identify negative sentiments accurately.

The Baseline Logistic Regression model with balanced classes (Other and Negative) is recommended for deployment.

It achieves the highest recall for the negative class, ensuring a higher number of negative sentiments are accurately identified, and performs quite well on the other classes ,providing the most balanced scores.

Next Steps

Deploy Selected Model:

- Implement the Baseline Logistic Regression model with balanced classes into a production environment.

Model Improvement:

- Establish a system to monitor the performance of the deployed model regularly.
- Collect feedback, track key metrics like precision, recall, and F1-score, and analyze any drifts in data or model accuracy.
- Schedule periodic retraining of the model with new, better labelled data to improve its performance and relevance.

Develop a Sentiment Response Strategy:

- Create a comprehensive strategy for responding to negative sentiments identified in the tweets.

Thank You !

Wambui Munene

wambui@icfoglobal.com