reagan-adajo / Phase_4

Code Issues Pull requests Actions Projects Wiki Security Insights Settings

0 stars 0 forks 1 watching Branches Activity
Tags

Public repository

master had recent pushes 59 seconds ago Compare & pull request

2 Branches 0 Tags Go to file t Go to file Add file Code

This branch is 6 commits ahead of, 1 commit behind `main`. Contribute

reagan-adajo revised commit 2a0f1af · 1 minute ago

.ipynb_checkpoints	third commit	3 days ago
README.md	revised commit	1 minute ago
image-1.png	fourth commit	3 days ago
image-2.png	fourth commit	3 days ago
image-3.png	fourth commit	3 days ago
image-4.png	revised commit	1 minute ago
image-5.png	revised commit	1 minute ago
image-6.png	revised commit	1 minute ago
image-7.png	revised commit	1 minute ago
image.png	fourth commit	3 days ago
index.ipynb	revised commit	1 minute ago
judge-1377884607_tweet_pro...	second commit	3 days ago
notebook.pdf	revised commit	1 minute ago
presentation.pdf	revised commit	1 minute ago
presentation.pptx	revised commit	1 minute ago

README

Phase IV Project Reagan Adajo

Summary: The project utilize two classification models to analyze the Tweets (Logistics Regression and Random Forest). The best Model will be determine for consequent adoption by the respective companies for automated sentiment rating

The outcomes of best-performing model will be used by the relevant companies to improve their brand, product, or service, among other decisions

Business and Data Understanding:

Business Problem

The company wants to develop a model that will assist in judging people's emotions about brands and products using Tweets on Apple and Google products.

Business Aim

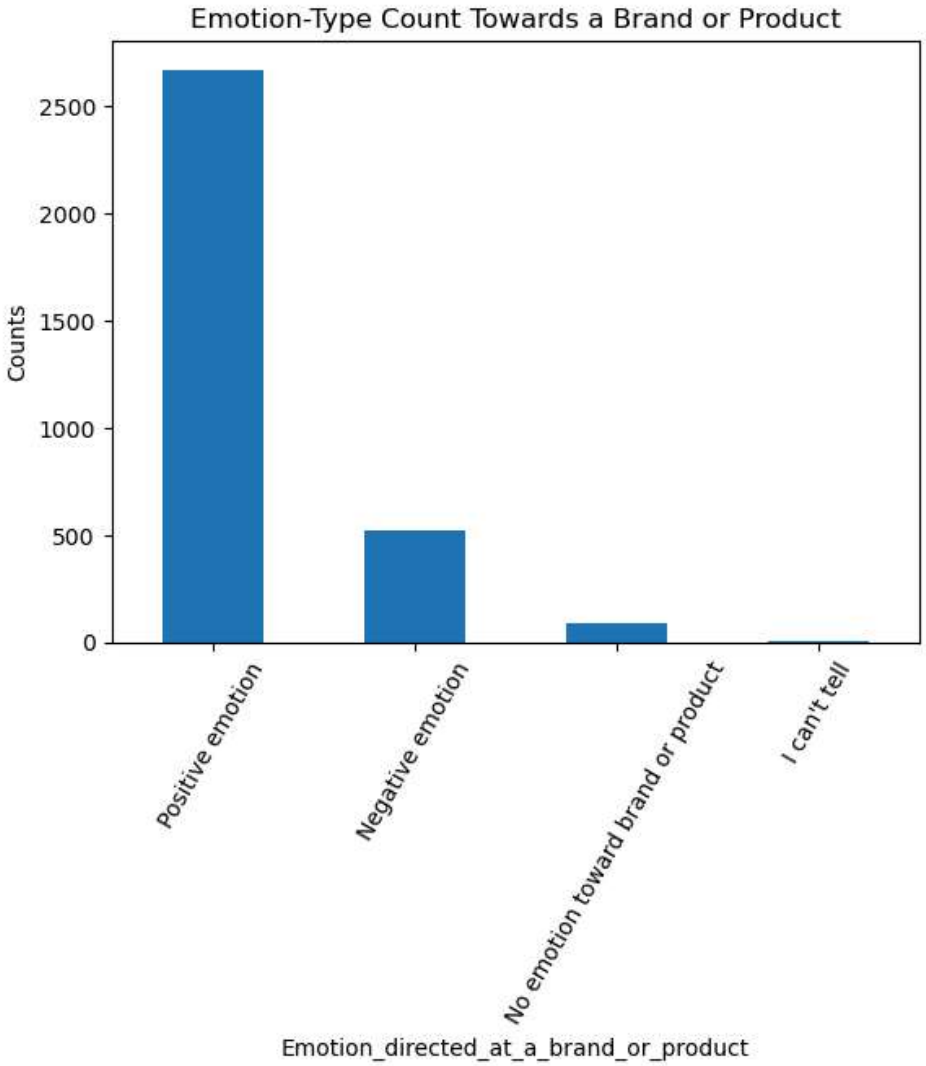
To build a model that judges customer's emotions on brands and products using their sentiments

Business Objectives

1. To find out if a Tweet can demonstrate a sentiment on Apple and Google products
2. To identify the products with the most positive and negative sentiments
3. To establish the best classification model for analyzing sentiments on Twitter texts(automatically)

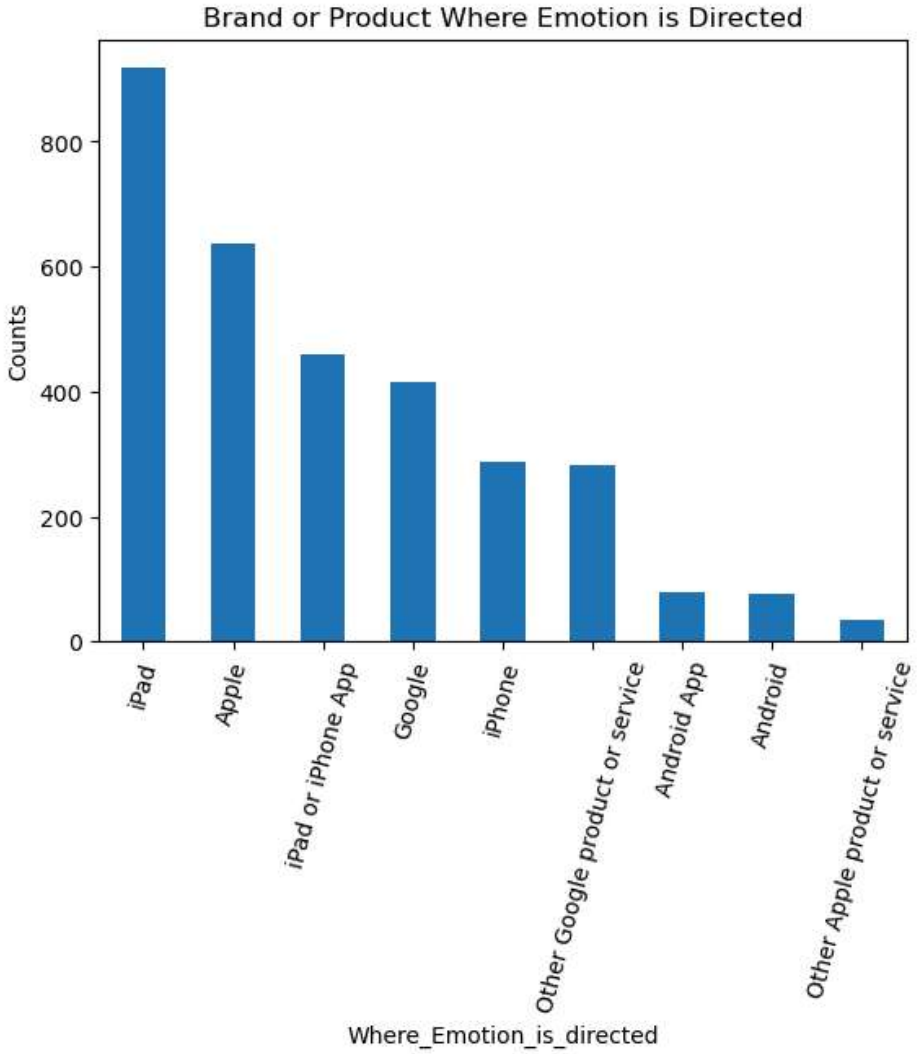
Data Understanding

The data was prepared by importing various python libraries along with various visualization to further explore



and understand the data.

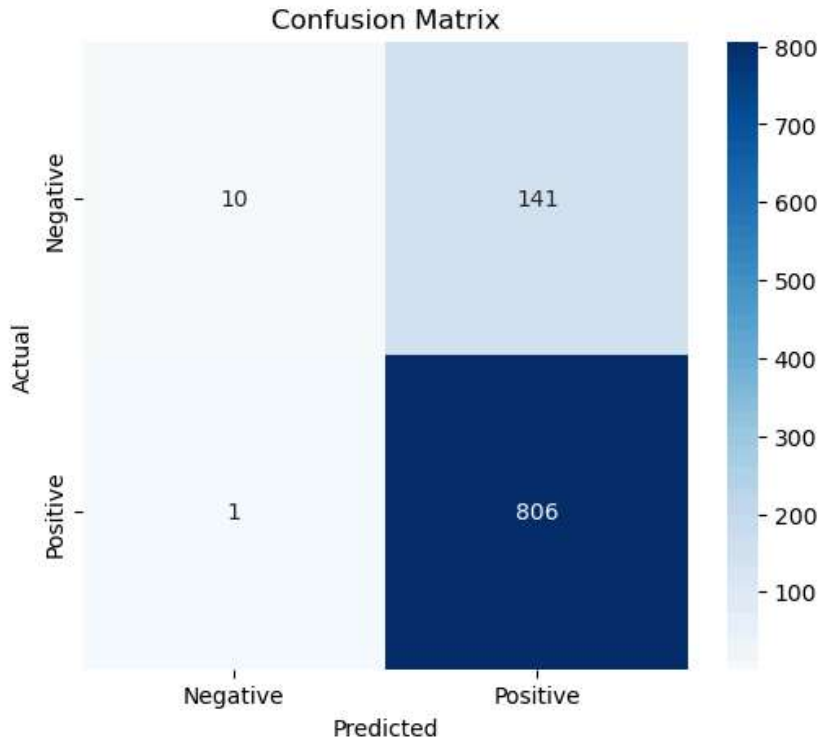
Modeling: A classification model was used on the data set to predict the where emotions were directed by the



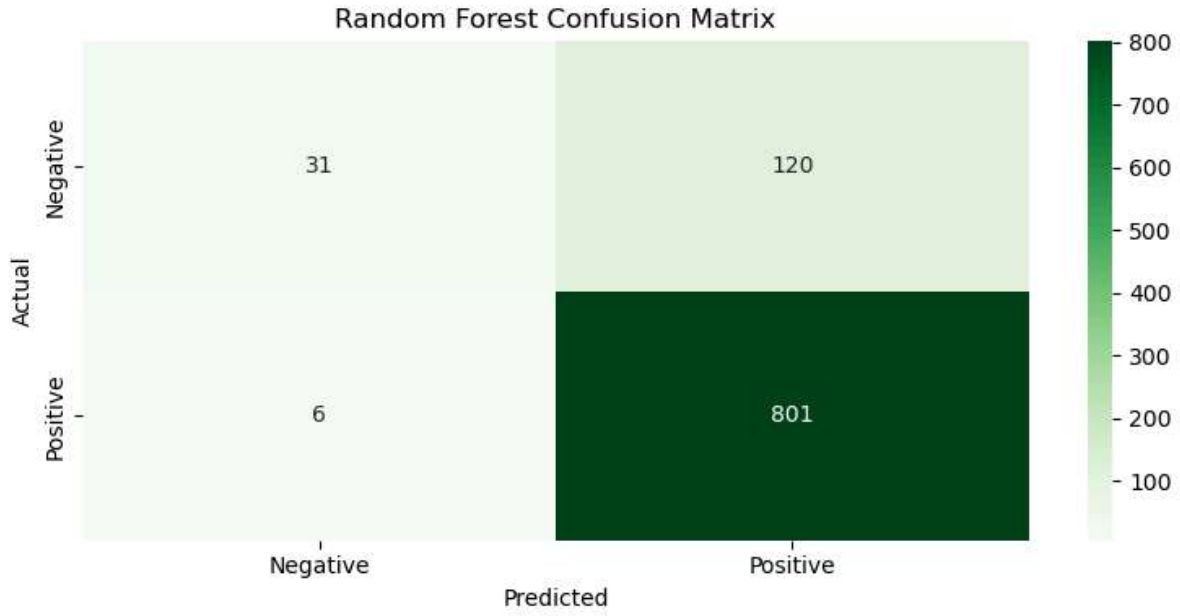
Tweets

The unique words were identified in the Tweets

Evaluation: The data was evaluated using Confusion Matrix. LogisticRegression !



RandomForest



Random Forest with Accuracy an accuracy of 0.8684759916492694 Conclusions: After performing EDA, preprocessing the text, and evaluating models, we conclude that Random Forest is the best with an accuracy of 0.8685.

Recommendation and Insights

Releases

No releases published
Create a new release

Packages

No packages published
Publish your first package