MORINGA

Discover · Grow · Transform

Group 7

Members:

Rodgers Bobb

Vitelis Siocha

George Odhiambo

Tweet Analysis on Google and Apple Products.

Positive

Negative

Nuetral

Google

Project Overview

- Project's objective is to develop a model to detect and analyse Twitter sentiment about Apple and Google products, by rating the sentiment as either positive, negative, or neutral basing on the content of the tweet.which is valuable when doing product analysis, rating, improving customer satisfaction as well as competitor analysis within companies.

Business understanding

- From the dataset at disposal, Google and Apple company intends to know the public perceptions on their products through analyzing tweet sentiments for further decision making, in order to bit the competitive dynamic market.

Google

Data Sources

- Data project at disposal is of csv format from the Data world, critical care needs to be checked before further analysis, techniques such as handling:
- Missing values
- Duplicates
- Tokenization
- Stemming
- Lemmatization
- Removal of digits and stop words.
- Part of Speech Tagging.

Data Preview

Dataset comprises 9,093 rows and 3 columns, with missing values present in the first two variables,

Data types is of object types, various needs to be deployed for further execution.

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 9093 entries, 0 to 9092
Data columns (total 3 columns):
 #   Column    Non-Null Count  Dtype 
---  ------    --------------  ----- 
 0   tweet     9092 non-null   object
 1   product   3291 non-null   object
 2   emotions  9093 non-null   object
dtypes: object(3)
memory usage: 213.2+ KB
```

	tweet_text	emotion_in_tweet_is_directed_at	is_there_an_emotion_directed_at_a_brand_or_product
0	.@wesley83 I have a 3G iPhone. After 3 hrs twe...	iPhone	Negative emotion
1	@jessedee Know about @fludapp ? Awesome iPad/i...	iPad or iPhone App	Positive emotion
2	@swonderlin Can not wait for #iPad 2 also. The...	iPad	Positive emotion
3	@sxsw I hope this year's festival isn't as cra...	iPad or iPhone App	Negative emotion
4	@sxtxstate great stuff on Fri #SXSW: Marissa M...	Google	Positive emotion
...	...	...	...
9088	Ipad everywhere. #SXSW {link}	iPad	Positive emotion
9089	Wave, buzz... RT @mention We interrupt your re...	NaN	No emotion toward brand or product
9090	Google's Zeiger, a physician never reported po...	NaN	No emotion toward brand or product
9091	Some Verizon iPhone customers complained their...	NaN	No emotion toward brand or product
9092	□Ï¡□Ïà□ü_□□Ê□□Î□□Ò□□£□□Á□âå□□_□□£□□□□å_□ÛâRT @...	NaN	No emotion toward brand or product

9093 rows × 3 columns

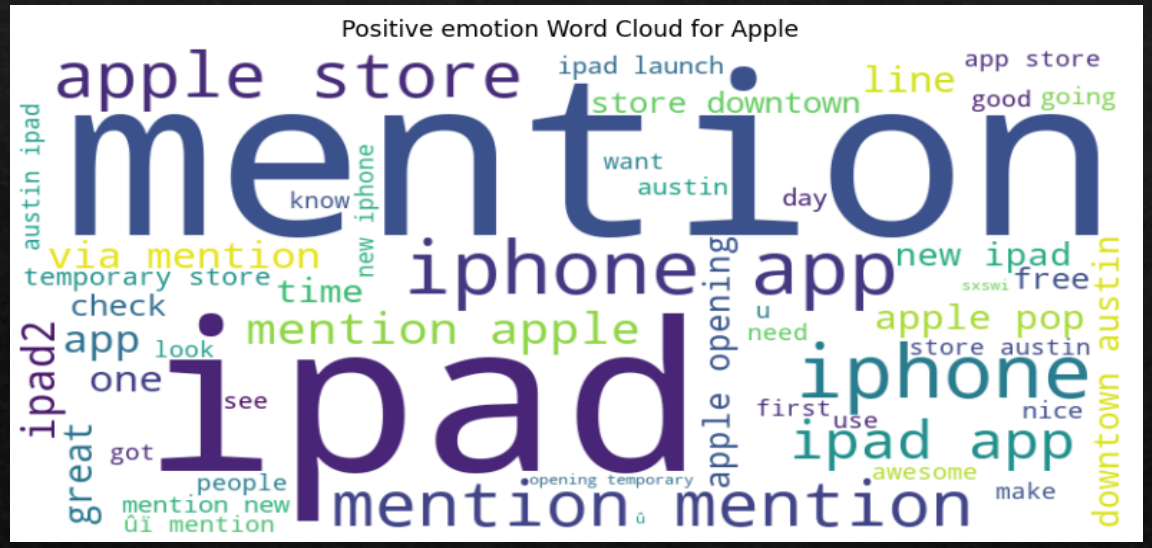
Missing Values

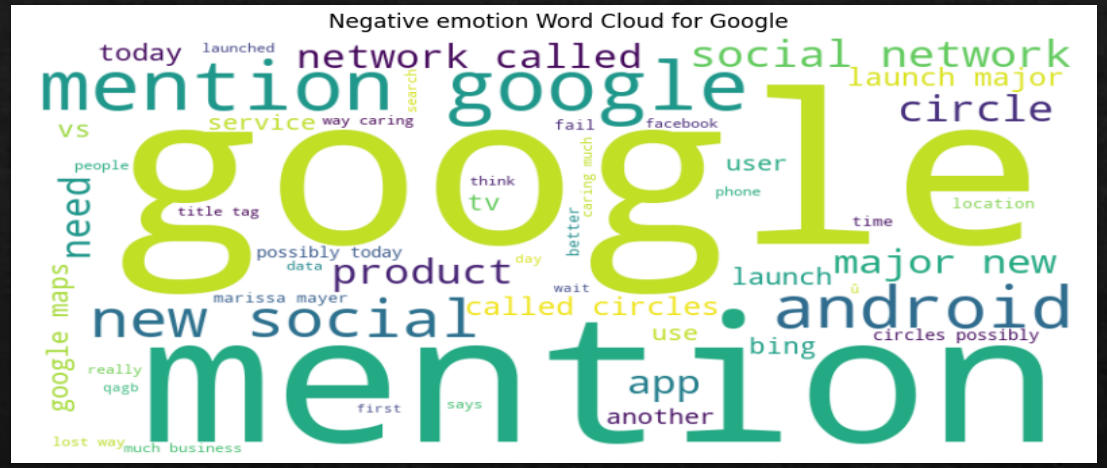
product	
NaN	5802
iPad	946
Apple	661
iPad or iPhone App	470
Google	430
iPhone	297
Other Google product or service	293
Android App	81
Android	78
Other Apple product or service	35

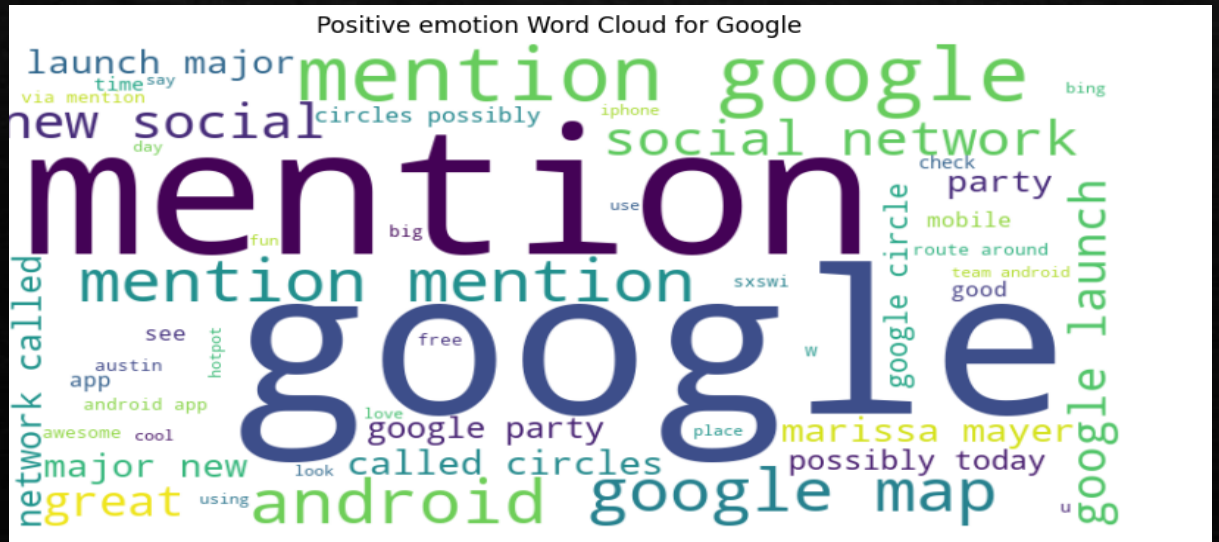
Name: count, dtype: int64

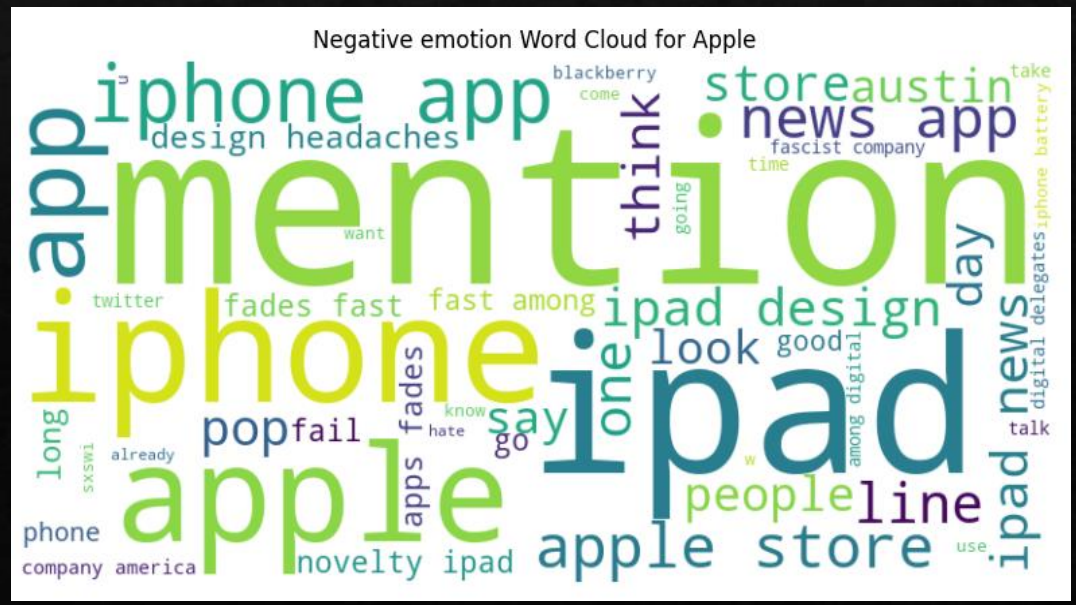
“Product" column has 5,802 missing (NaN), with other categories such as iPad, Apple, and Google also present. Notably, similar products are categorized differently, e.g, Android App and Android.

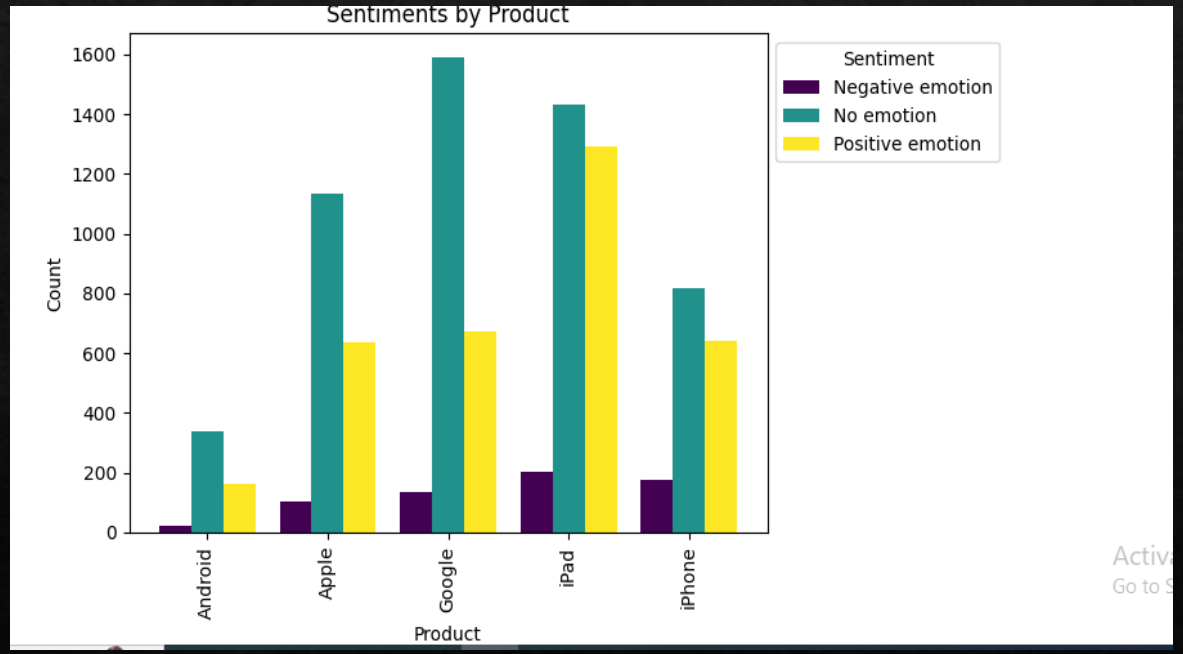
Sentiments Visualizations

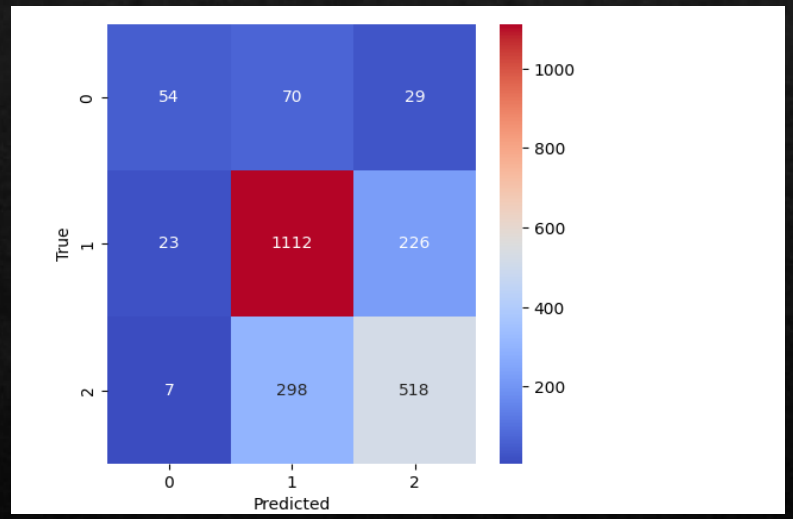




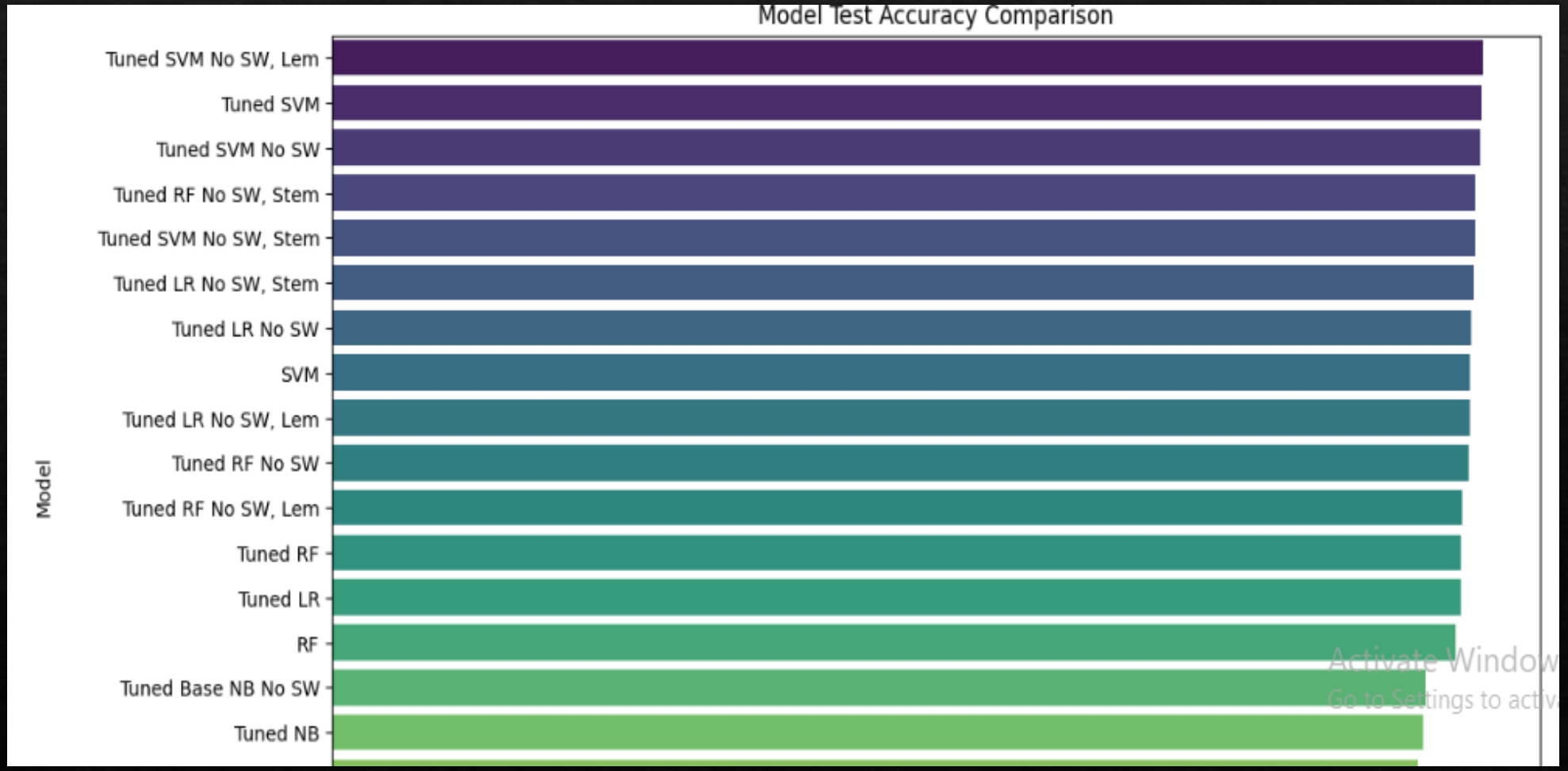





Sentiments by Product
Sentiment
Negative emotion
No emotion
Positive emotion
Count
0
200
400
600
800
1000
1200
1400
1600
Android
Apple
Google
iPad
iPhone
Product


54
70
29
23
1112
226
7
298
518
True
Predicted
0
1
2
200
400
600
800
1000

Optimal Model Evaluation



Conclusions

- Sentiments pictorials (wordcloud and bar graph) discussions revolves around Apple and Google, with users expressing both positive, neutral and negative sentiments towards the companies and their products. This indicates diverse opinions and engagement across platforms about Google and Apple products.
- The optimized Support Vector Machines (SVM) model, built from lemmatized tweets without stop words, emerged as the best performer with a 72.9% accuracy of classifying unseen tweets based on sentiments.

Limitations

- Dataset at disposal is small,
- Only one platform of data source was given which is twitter.
- Opinions are personal, and the model might not understand all the different ways people express themselves. Sarcasm and subtle language can be tricky for the tool to understand. Significant class imbalance, with a majority of tweets being neutral (61%), poses challenges in accurately capturing and addressing negative sentiments, potentially affecting the identification of areas for improvement.

Recommendations

- 1 Enhance Product Features Based on the tweet sentiments: Apple and Google should proactively address user concerns and criticisms.
- 2. Diversify Data Sources for a Comprehensive View: Expand data collection beyond Twitter to include opinions from specific groups or smaller markets not covered by the current dataset, ensuring a more comprehensive understanding.
- 3. Implement Sentiment Analysis Refinement: Explore the integration of Neural Networks to significantly enhance model performance. Consider incorporating Deep Learning methodologies in future projects for more accurate and enhanced tweet classification.
- 4. Strategic Marketing for Positive Sentiments: Utilize the buzz around Google's "Circle" project as a valuable marketing tool. Integrate it strategically into promotional efforts to magnify positive sentiments, encouraging greater brand interaction and conversation. This strategic approach enhances the overall public image, potentially attracting a larger customer base and strengthening Google's market standing.
- 5. Address Neutral Sentiments Through Proactive Measures to make them positive: Despite the minimal negative sentiments, proactively address any identified areas for improvement to maintain a positive brand image and stay competitive in the market. Also, conduct a focused analysis to discern sentiments of the neutral majority, allowing for focused strategies that can enhance both brand perception and market promotion. This can also help to address the unreavealed negative sentiments existing amongst the neutral majority class.

Thank You


NLP
Natural language processing
shutterstock.com · 1467529412