SENTIMENT ANALYSIS FOR MARKETING

TEAM MEMBER

NAME:S.SANMATHI

REGNO: 912421106305

Gmail id: ssanmathi106@gmail.com

NM ID: aut2291240016

PHASE 2 Submission Document

Introduction

- Sentiment analysis can be defined as analyzing the positive or negative sentiment of the customer in text. The contextual analysis of identifying information helps businesses understand their customers' social sentiment by monitoring online conversations.
- As customers express their reviews and thoughts about the brand more openly than ever before, sentiment analysis has become a powerful tool to monitor and understand online conversations.
- Recent advancements in machine learning and deep learning have increased the efficiency of sentiment analysis algorithms. You can creatively use advanced artificial intelligence and machine learning tools for doing research and draw out the analysis.

Content for Project Phase 2

Consider exploring advanced techniques like fine-tuning pre-trained sentiment analysis models (BERT, RoBERTa) more accurate sentiment predictions.

Data source

Sentiment analysis on customer feedback to gain insights into competitor products. By understanding customer sentiments, companies can identify strengths and weaknesses in competing products, thereby improving their own offerings. This project requires utilizing various NLP methods to extract valuable insights from customer feedback.

Dataset Link

https://www.kaggle.com/datasets/crowdflower/twitter-airline-sentiment

tweet_id	airline_se	airline_se	negativer	negativer	airline	airline_se	name	negativer	retweet_c	text	tweet_coo	tweet_cre	tweet_loc	user_timezone
5.7E+17	neutral	1			Virgin America		cairdin		0	@VirginAmerica Wh		########		Eastern Time (US & Canada)
5.7E+17	positive	0.3486		0	Virgin America		jnardino		0	@VirginAmerica plu		########		Pacific Time (US & Canada)
5.7E+17	neutral	0.6837			Virgin America		yvonnalynn		0	@VirginAmerica I di		########	Lets Play	Central Time (US & Canada)
5.7E+17	negative	1	Bad Flight	0.7033	Virgin America		jnardino		0	@VirginAmerica it's		########		Pacific Time (US & Canada)
5.7E+17	negative	1	Can't Tell	1	Virgin America		jnardino		0	@VirginAmerica and		########		Pacific Time (US & Canada)
5.7E+17	negative	1	Can't Tell	0.6842	Virgin America		jnardino		0	@VirginA		########		Pacific Time (US & Canada)
5.7E+17	positive	0.6745		0	Virgin America		cjmcginnis		0	@VirginAmerica yes		########	San Franci	Pacific Time (US & Canada)
5.7E+17	neutral	0.634			Virgin America		pilot		0	@VirginAmerica Rea		########	Los Angel	Pacific Time (US & Canada)
5.7E+17	positive	0.6559			Virgin America		dhepburn		0	@virginamerica Wel		########	San Diego	Pacific Time (US & Canada)
5.7E+17	positive	1			Virgin America		YupitsTate		0	@VirginAmerica it w		########	Los Angel	Eastern Time (US & Canada)
5.7E+17	neutral	0.6769		0	Virgin America		idk_but_youtube		0	@VirginAmerica did		########	1/1 loner	Eastern Time (US & Canada)
5.7E+17	positive	1			Virgin America		HyperCamiLax		0	@VirginAmerica I &I		########	NYC	America/New_York
5.7E+17	positive	1			Virgin America		HyperCamiLax		0	@VirginAmerica This		########	NYC	America/New_York
5.7E+17	positive	0.6451			Virgin America		mollanderson		0	@VirginAmerica @v		########		Eastern Time (US & Canada)
5.7E+17	positive	1			Virgin America		sjespers		0	@VirginAmerica Tha		########	San Franci	Pacific Time (US & Canada)
5.7E+17	negative	0.6842	Late Flight	0.3684	Virgin America		smartwatermelon		0	@VirginAmerica SFC		########	palo alto,	Pacific Time (US & Canada)
5.7E+17	positive	1			Virgin America		ItzBrianHunty		0	@VirginAmerica So e		########	west covir	Pacific Time (US & Canada)
5.7E+17	negative	1	Bad Flight	1	Virgin America		heatherovieda		0	@VirginAmerica I fl		########	this place	Eastern Time (US & Canada)
5.7E+17	positive	1			Virgin America		thebrandiray		0	I â[illegible]flying @VirginA		########	Somewhe	Atlantic Time (Canada)
5.7E+17	positive	1			Virgin America		JNLpierce		0	@VirginAmerica you		########	Boston | V	Quito
5.7E+17	negative	0.6705	Can't Tell	0.3614	Virgin America		MISSGJ		0	@VirginAmerica why		########		
5.7E+17	positive	1			Virgin America		DT_Les		0	@VirginA	[40.74804	########		
5.7E+17	positive	1			Virgin America		ElvinaBeck		0	@VirginAmerica I lo		########	Los Angel	Pacific Time (US & Canada)
5.7E+17	neutral	1			Virgin America		rjlynch21086		0	@VirginAmerica will		########	Boston, M	Eastern Time (US & Canada)
5.7E+17	negative	1	Customer	0.3557	Virgin America		ayeevickiee		0	@VirginAmerica you		########	714	Mountain Time (US & Canada)
5.7E+17	negative	1	Customer	1	Virgin America		Leora13		0	@VirginAmerica stat		########		
5.7E+17	negative	1	Can't Tell	0.6614	Virgin America		meredithjlynn		0	@VirginAmerica Wh		########		
5.7E+17	neutral	0.6854			Virgin America		AdamSinger		0	@VirginAmerica do		########	San Franci	Central Time (US & Canada)
5.7E+17	negative	1	Bad Flight	1	Virgin America		blackjackpro911		0	@VirginA	[42.36101	########	San Mateo, CA & Las Vegas, NV	
5.7E+17	neutral	0.615		0	Virgin America		TenantsUpstairs		0	@VirginA	[33.94540	########	Brooklyn	Atlantic Time (Canada)
5.7E+17	negative	1	Flight Boo	1	Virgin America		jordanpichler		0	@VirginAmerica hi!		########		Vienna
5.7E+17	neutral	1			Virgin America		JCervantezzz		0	@VirginAmerica Are		########	California	Pacific Time (US & Canada)
5.7E+17	negative	1	Customer	1	Virgin America		Cuschoolie1		0	@VirginA	[33.94209	########	Washingt	Quito
5.7E+17	negative	1	Customer	1	Virgin America		amanduhmccarty		0	@VirginAmerica awa		########		Pacific Time (US & Canada)

Data Collection:

Identify a dataset containing customer reviews and sentiment about competitor products.

Data Preprocessing:

Clean and preprocess the textual data for analysis.

Sentiment analysis techniques:

Employ different NLP techniques like Bag of words, word embeddings, or transformer models for sentiment analysis.

Feature extraction:

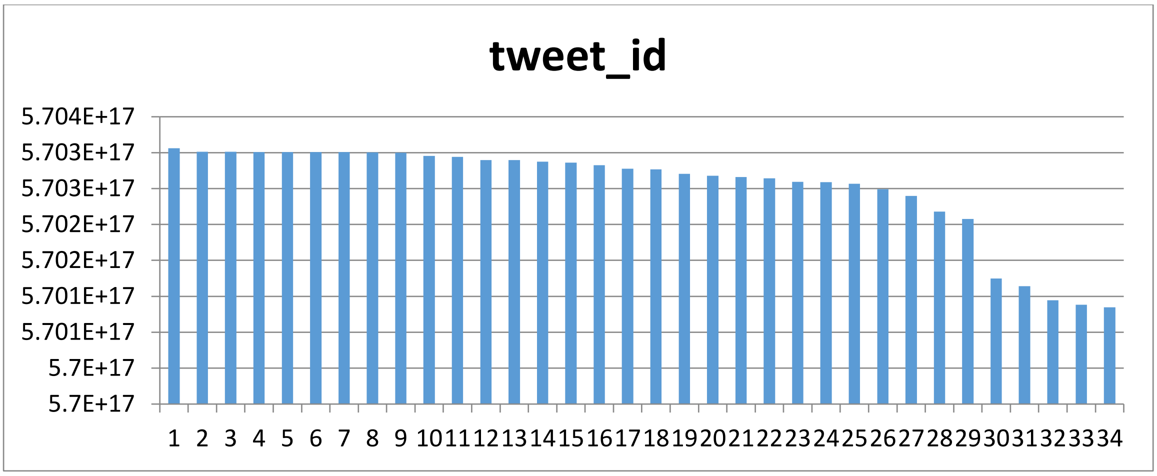
Extract features and sentiment from the text data

Visualization:

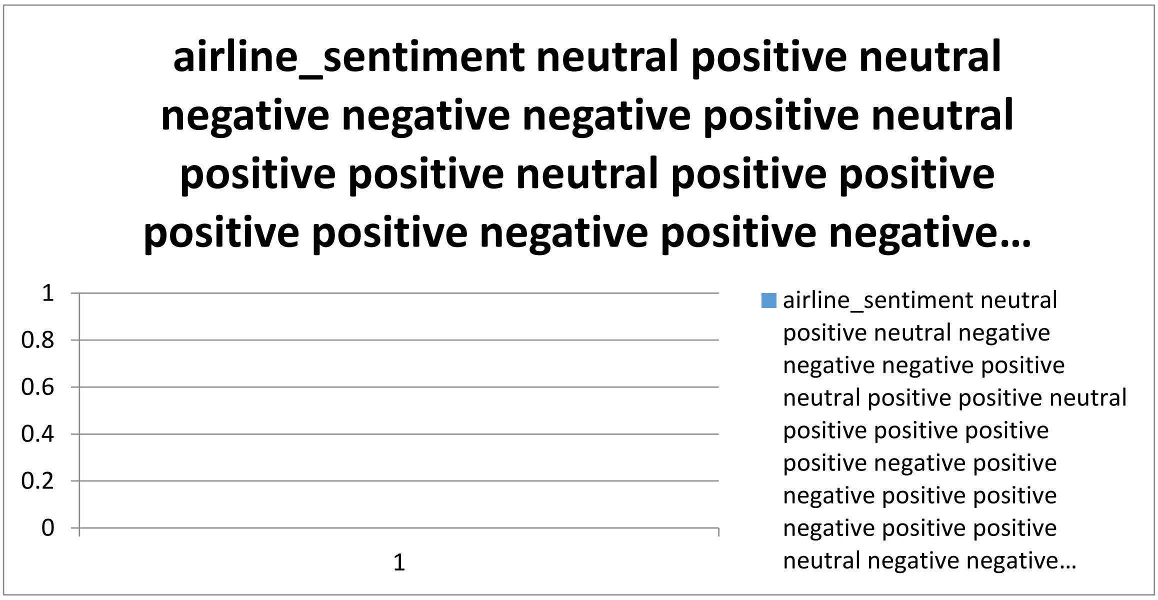
Create visualization to depict the sentiment distribution and analyze trends.

Insights Generation:

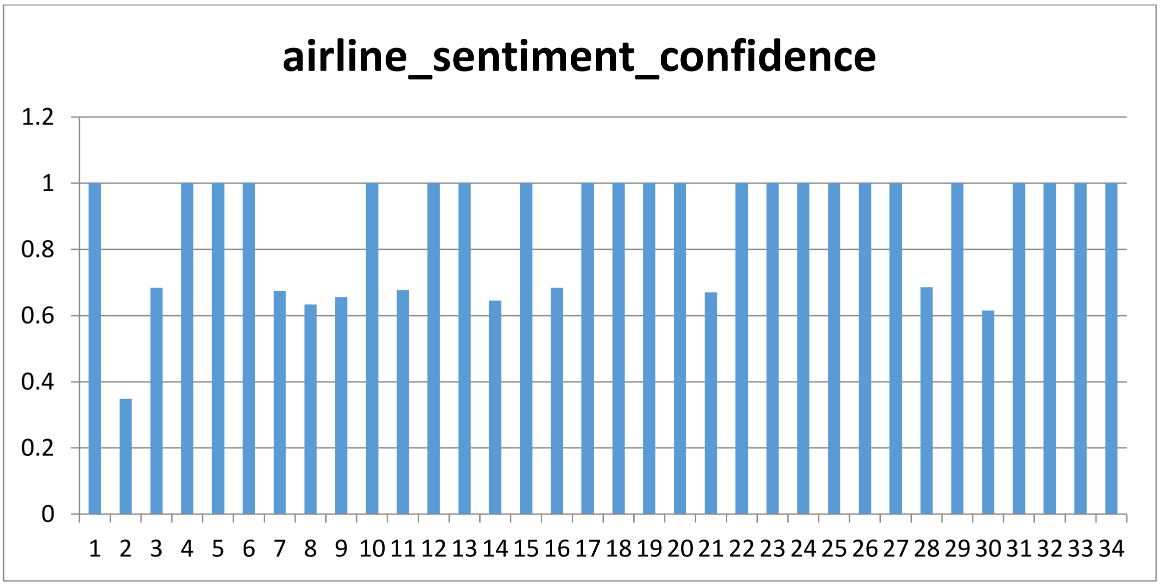
Extract meaningful insights from the sentiment analysis result to guide business decisions.



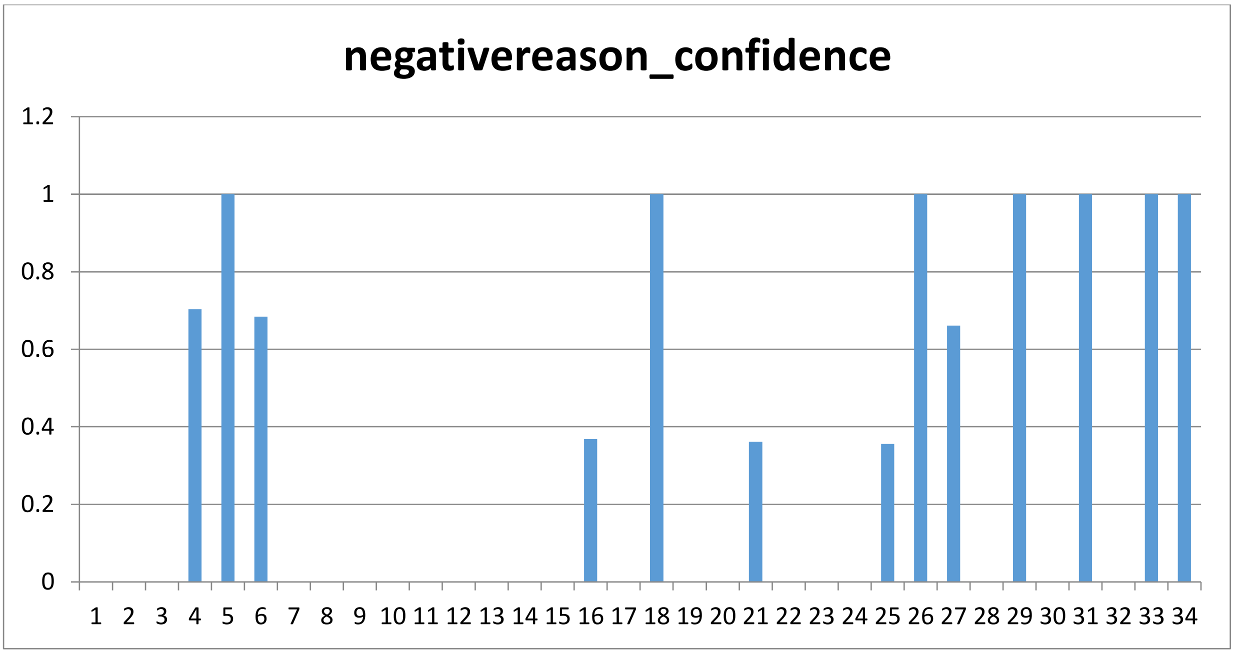
tweet_id
5.704E+17
5.703E+17
5.703E+17
5.702E+17
5.702E+17
5.701E+17
5.701E+17
5.7E+17
5.7E+17
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34

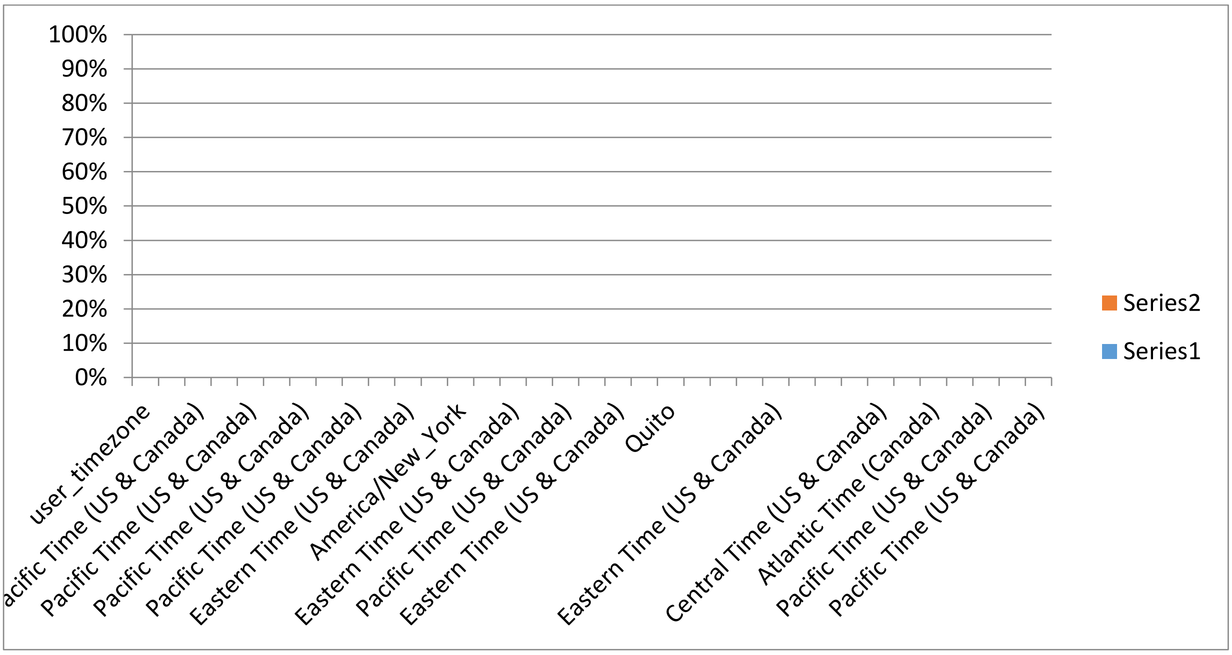


airline_sentiment neutral positive neutral negative negative negative positive neutral positive positive neutral positive positive positive positive negative positive negative...
1
0.8
0.6
0.4
0.2
0
1
airline_sentiment neutral positive neutral negative negative negative positive neutral positive positive neutral positive positive positive positive negative positive negative positive positive negative positive positive neutral negative negative...



airline_sentiment_confidence
1.2
1
0.8
0.6
0.4
0.2
0
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34


negativereason_confidence
1.2
1
0.8
0.6
0.4
0.2
0
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34


100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
user_timezone
acific Time (US & Canada)
Pacific Time (US & Canada)
Pacific Time (US & Canada)
Pacific Time (US & Canada)
Eastern Time (US & Canada)
America/New_York
Eastern Time (US & Canada)
Pacific Time (US & Canada)
Eastern Time (US & Canada)
Quito
Eastern Time (US & Canada)
Central Time (US & Canada)
Atlantic Time (Canada)
Pacific Time (US & Canada)
Pacific Time (US & Canada)
Series2
Series1

Types of Sentiment Analysis

Various types of sentiment analysis can be performed, depending on the specific focus and objective of the analysis. Some common types include:

- **Document-Level Sentiment Analysis**: This type of analysis determines the overall sentiment expressed in a document, such as a review or an article. It aims to classify the entire text as positive, negative, or neutral.

- **Sentence-Level Sentiment Analysis:** Here, the sentiment of each sentence within a document is analyzed. This type provides a more granular understanding of the sentiment expressed in different text parts.

- **Aspect-Based Sentiment Analysis:** This approach focuses on identifying and extracting the sentiment associated with specific aspects or entities mentioned in the text. For example, in a product review, the sentiment towards different features of the product (e.g., performance, design, usability) can be analyzed separately.

- **Entity-Level Sentiment Analysis:** This type of analysis identifies the sentiment expressed towards specific entities or targets mentioned in the understand the sentiment associated with different entities within the same document.

- **Comparative Sentiment Analysis:** This approach involves comparing the sentiment between different entities or aspects mentioned in the text. It aims to identify the relative sentiment or preferences expressed towards various entities or features.

Sentiment Analysis Use Cases

We just saw how sentiment analysis can empower organizations with insights that can help them make data-driven decisions.

Social Media Monitoring for Brand Management: Brands can use sentiment analysis to gauge their Brand's public outlook.

Product/Service Analysis: Brands/Organizations can perform sentiment analysis on customer reviews to see how well a product or service is doing in the market and make future decisions accordingly.

Stock Price Prediction: Predicting whether the stocks of a company will go up or down is crucial for investors.

Program

SENTIMENT ANALYSIS FOR MARKETING

Code for Sentiment Analysis Using Vader:

```
from vaderSentiment.vaderSentiment import SentimentIntensityAnalyzer

sentiment = SentimentIntensityAnalyzer()

text_1 = "The book was a perfect balance between wrtiting style and plot."

text_2 =  "The pizza tastes terrible."

sent_1 = sentiment.polarity_scores(text_1)

sent_2 = sentiment.polarity_scores(text_2)

print("Sentiment of text 1:", sent_1)

print("Sentiment of text 2:", sent_2)
```

Output

```
Sentiment of text 1: {'neg': 0.0, 'neu': 0.73, 'pos': 0.27, 'compound': 0.5719}

Sentiment of text 2: {'neg': 0.508, 'neu': 0.492, 'pos': 0.0, 'compound': -0.4767}
```

Code for Sentiment Analysis using Bag of Words Vectorization Approach:

```
#Loading the Dataset
import pandas as pd
data = pd.read_csv('Finance_data.csv')
#Pre-Prcoessing and Bag of Word Vectorization using Count Vectorizer
from sklearn.feature_extraction.text import CountVectorizer
from nltk.tokenize import RegexpTokenizer
token = RegexpTokenizer(r'[a-zA-Z0-9]+')
cv = CountVectorizer(stop_words='english',ngram_range = (1,1),tokenizer = token.tokenize)
text_counts = cv.fit_transform(data['sentences'])
#Splitting the data into trainig and testing
from sklearn.model_selection import train_test_split
X_train, X_test, Y_train, Y_test = train_test_split(text_counts, data['feedback'], test_size=0.25, random_state=5)
#Training the model
from sklearn.naive_bayes import MultinomialNB
MNB = MultinomialNB()
MNB.fit(X_train, Y_train)
#Caluclating the accuracy score of the model
from sklearn import metrics
```

```
predicted = MNB.predict(X_test)
accuracy_score = metrics.accuracy_score(predicted, Y_test)
print("Accuracuy Score: ",accuracy_score)
```

Output:

Accuracy Score: 0.9111675126903553

Code for Sentiment Analysis Using Transformer based models:

```
from transformers import pipeline
sentiment_pipeline = pipeline("sentiment-analysis")
data = ["It was the best of times.", "t was the worst of times."]
sentiment_pipeline(data)
```

Output:

[{'label': 'POSITIVE', 'score': 0.999457061290741}, {'label': 'NEGATIVE', 'score': 0.9987301230430603}]

Advantages of Sentiment Analysis

- product review monitoring – monitoring which of your products receive a higher rate of positive comments.
- market research – discovering attitudes of internet users toward the research target.
- search engines/recommender systems – enhancing performance by better understanding what users meant by the content of a query.

Conclusion

Sentiment analysis can be a very useful tool for user response monitoring. Its most significant advantage is the introduction of the possibility to use direct user feedback with minimal human supervision while still being able to scale up easily.