SENTIMENT ANALYSIS FOR MARKETING

BATCH MEMBER

NAME: V.SRINITHI

REGNO: 912421106307

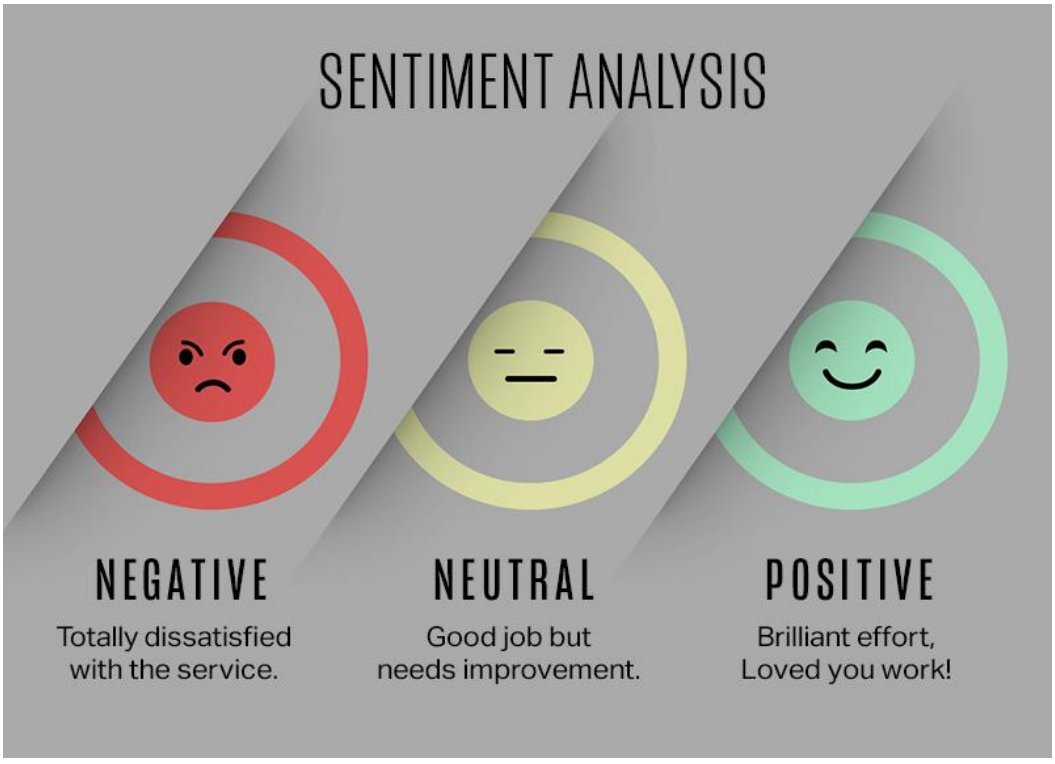
Gmail id: srinithiveera@gmail.com

NM ID: aut2291240021

PHASE 3 Submission Document

Phase 3: Development Part 1

Topic: Sentiment analysis for marketingbyloading and pre-processing the dataset



Introduction

- Sentiment analysis can be defined as analysing 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.

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 customerfeedback.

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

tweet id	airline_ser	airline_ser	negativere	negativere	airline	airline_ser	name	negativere
5.7E+17	negative	1	Customer	1	Virgin America		amanduhmccarty	
5.7E+17	positive	1			Virgin America		NorthTxHomeTeam	
5.7E+17	neutral	0.6207			Virgin America		miaerolinea	
5.7E+17	positive	1			Virgin America		Nicsplace	
5.7E+17	positive	1			Virgin America		Nicsplace	
5.7E+17	neutral	0.6791		0	Virgin America		elisha_malulani	
5.7E+17	negative	1	Customer	1	Virgin America		DannyDouglass	
5.7E+17	positive	0.6639			Virgin America		jamesferrandini	
5.7E+17	negative	0.6688	Flight Bool	0.6688	Virgin America		will_lenzenjr	
5.7E+17	neutral	1			Virgin America		GottAmanda	
5.7E+17	neutral	0.6578		0	Virgin America		KGervaise	
5.7E+17	neutral	1			Virgin America		papamurat	
5.7E+17	positive	1			Virgin America		arieldaie	
5.7E+17	neutral	0.6799			Virgin America		vacations7	
5.7E+17	positive	1			Virgin America		ChelseaPoe666	
5.7E+17	neutral	1			Virgin America		BobGlavinVO	
5.7E+17	neutral	0.6436			Virgin America		lisaaiko	
5.7E+17	neutral	0.6764		0	Virgin America		grantbrowne	
5.7E+17	positive	0.657			Virgin America		joyabsalon	
5.7E+17	neutral	1			Virgin America		2v	
5.7E+17	neutral	0.7118		0	Virgin America		KSmithFoundHere	
5.7E+17	neutral	1			Virgin America		papamurat	
5.7E+17	negative	0.6939	Flight Bool	0.6939	Virgin America		murphicus	
5.7E+17	positive	1			Virgin America		VinnieFerra	
5.7E+17	positive	0.635			Virgin America		KevinDemsi	
5.7E+17	neutral	0.7007			Virgin America		giffgaffman	
5.7E+17	neutral	1			Virgin America		HanlonBrothers	
5.7E+17	neutral	1			Virgin America		emilybg78	
5.7E+17	negative	1	Customer	1	Virgin America		rachie1126	
5.7E+17	neutral	0.6858			Virgin America		adawson66	
5.7E+17	neutral	1			Virgin America		SocialPLC	
5.7E+17	positive	1			Virgin America		jeffreymace01	
5.7E+17	neutral	0.6814		0	Virgin America		1stcrown	
5.7E+17	negative	1	Customer	1	Virgin America		onerockgypsy	
5.7E+17	negative	1	Late Flight	0.6789	Virgin America		noelduan	
5.7E+17	positive	0.6922			Virgin America		Travelzoo	

Data Collection:

Identify a dataset containing customer reviews and sentiment aboutcompetitor products.

Data Preprocessing:

Clean and preprocess the textual data for analysis.

- Step 1: Delete duplicate data.
- Step 2: Remove irrelevant items.
- Step 3: Check for outlier data.
- Step 4: Correct typos and structural mistakes.
- Step 5: Check for missing data.
- Step 6: Validate your data.
- Discover More: Complete Sentiment Analysis Process.
- The data pre-processing techniques includes five activities such as Data Cleaning, Data Optimization, Data Transformation, Data Integration and Data Conversion.

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:

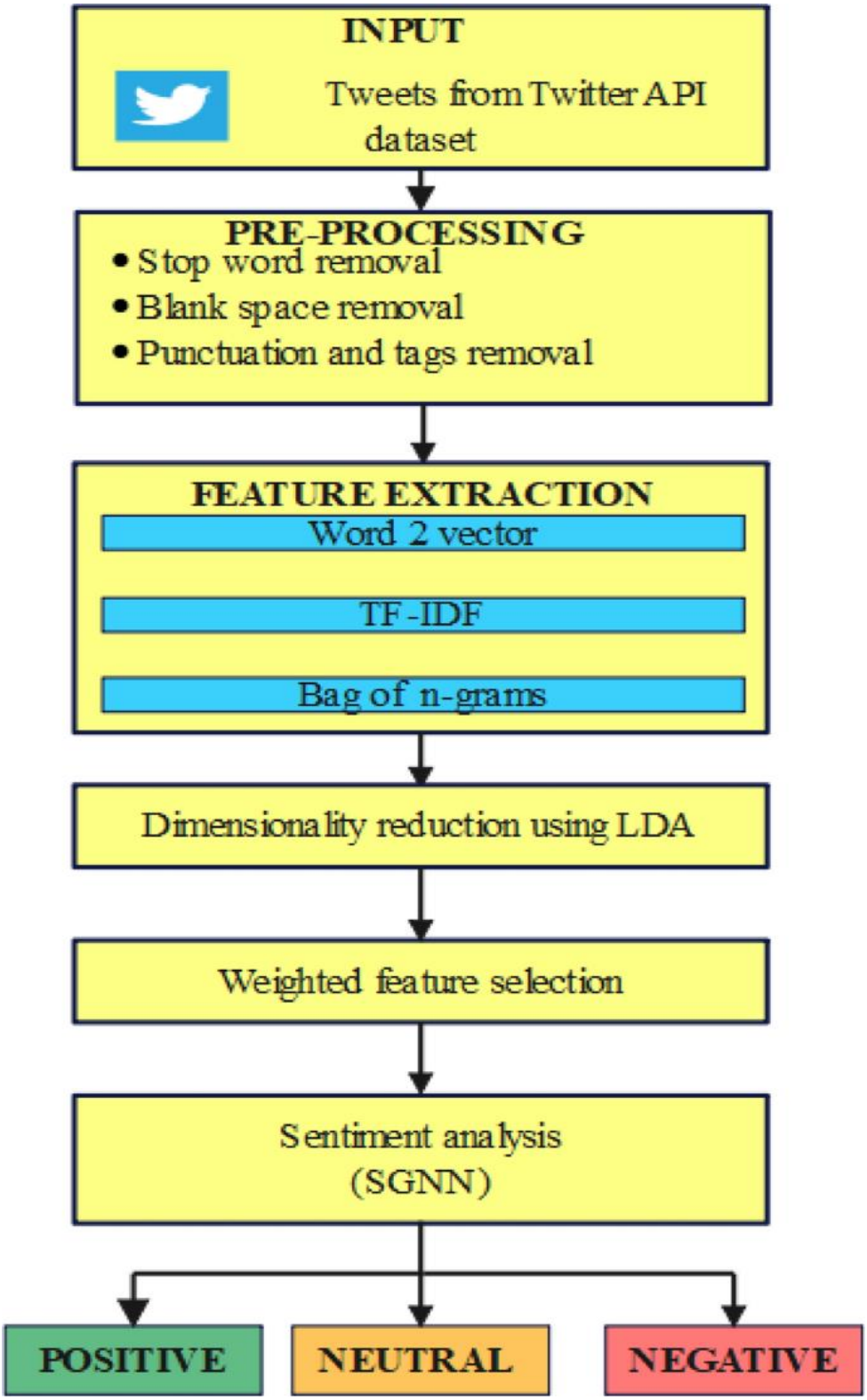
Extract meaningful insights from the sentiment analysis result to guide business decisions.

Sentiment analysis techniques:

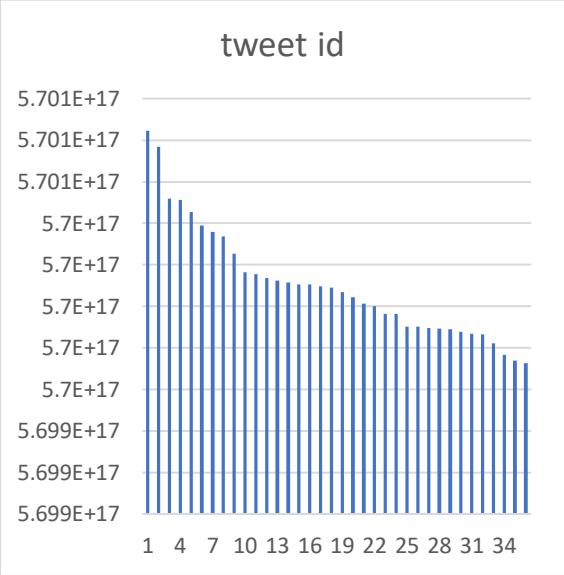
- NLP model:A rule-based system uses a set of human-crafted rules to help identify subjectivity, polarity, or the subject of an opinion.
- These rules may include various NLP techniques developed in computational linguistics, such as: Stemming, tokenization, part-of-speech tagging and parsing.
- bag-of-words model in sentiment analysis: bag-of-words model is a way of extracting features from text so the text input can be used with machine learning algorithms like neural networks.
- Each document, in this case a review, is converted into a vector representation.

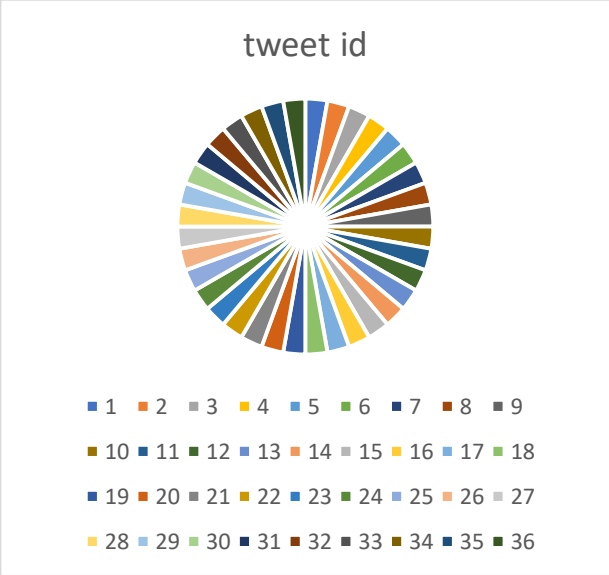
Proposed system:

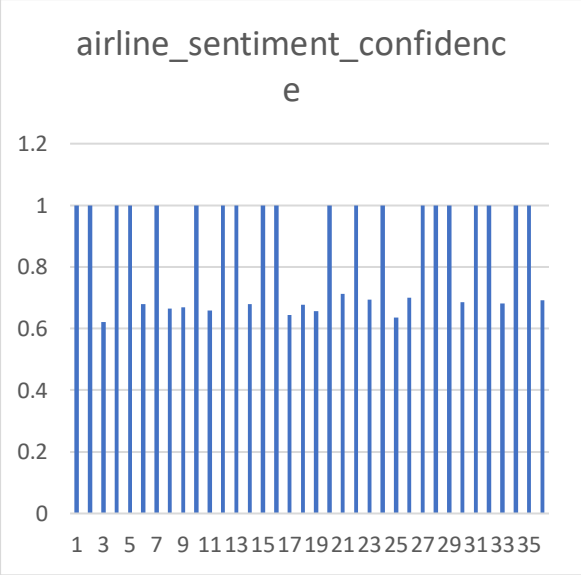
- Performing sentiment analysis on social networks such as Twitter is considered as a significant tool to gather information about the opinion or emotions of the public in real-time applications.
- The proposed work helps in extracting the sentiment of the tweet posted by Twitter users in various situations and the proposed method have the capability to recognize the emotion from the text.
- The machine learning method is used for analysing the sentiment which helps to gain the capability automatic learning to the model. The block diagram for sentiment analysis using in Twitter data using machine learning method.

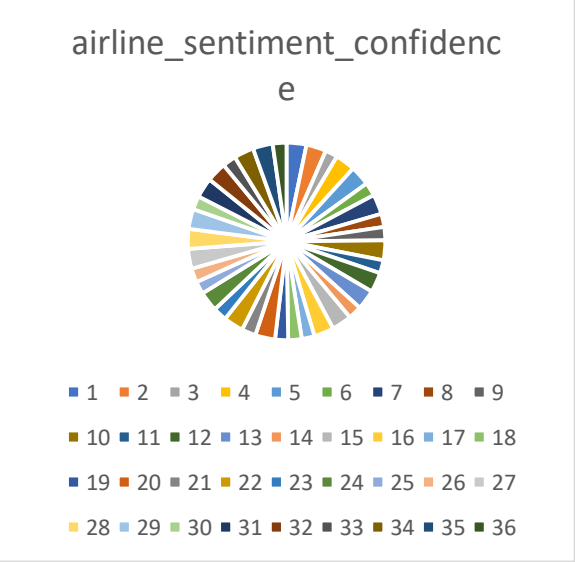

INPUT
Tweets from Twitter API dataset
PRE-PROCESSING
• Stop word removal
• Blank space removal
• Punctuation and tags removal
FEATURE EXTRACTION
Word 2 vector
TF-IDF
Bag of n-grams
Dimensionality reduction using LDA
Weighted feature selection
Sentiment analysis (SGNN)
POSITIVE
NEUTRAL
NEGATIVE

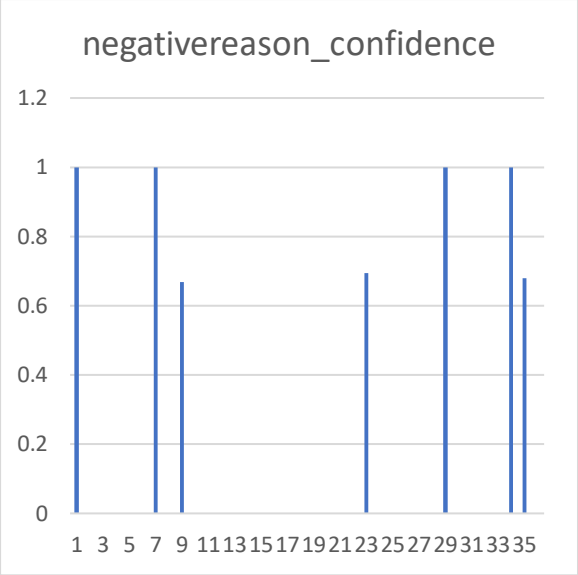
Pre-Processing data set:

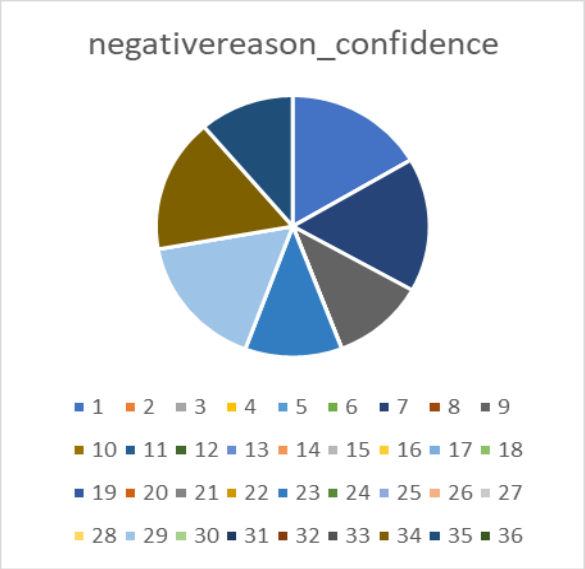


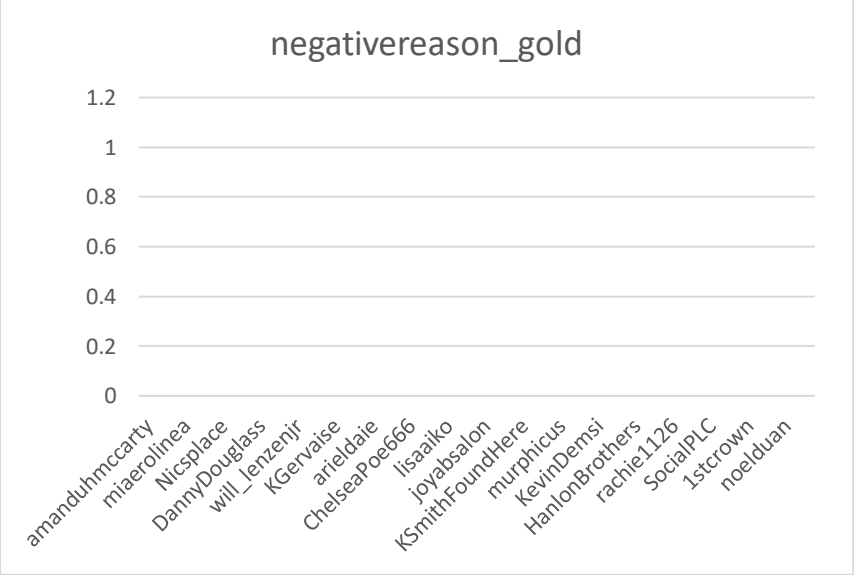


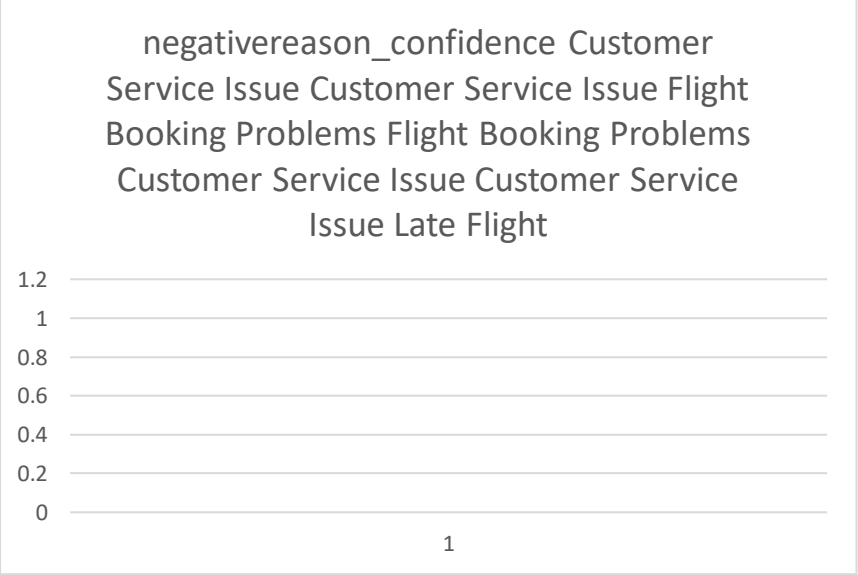


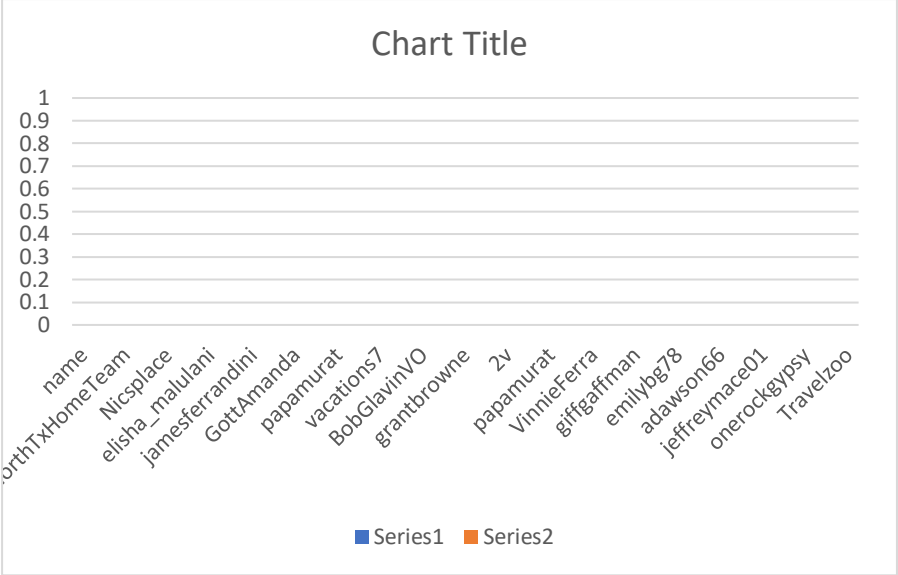



negativereason_confidence
1.2
1
0.8
0.6
0.4
0.2
0
1 3 5 7 9 11 13 15 17 19 21 23 25 27 29 31 33 35


negativereason_confidence
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 35 36


negativereason_gold
1.2
1
0.8
0.6
0.4
0.2
0
amanduhmccarty
miaerolinea
Nicsplace
DannyDouglass
will_lenzenjr
KGervaise
arieldaie
ChelseaPoe666
lisaalko
joyabsalon
KSmithFoundHere
murphicus
KevinDemsi
HanlonBrothers
rachie1126
SocialPLC
1stcrown
noelduan


negativereason_confidence Customer Service Issue Customer Service Issue Flight Booking Problems Flight Booking Problems Customer Service Issue Customer Service Issue Late Flight
1.2
1
0.8
0.6
0.4
0.2
0
1



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.