Experiment No.6
Text analytics: Implementation of Spam filter/Sentiment analysis in python/R.
Date of Performance:
Date of Submission:

EXPERIMENT NO 6

Aim: Implementation of Sentiment Analysis

Objective: To understand the use of various sentiment Analysis techniques by implementing Them.

Description:

Sentiment Analysis: Sentiment Analysis is a text analysis technique that allows companies to make sense of qualitative data. By detecting positive and negative sentiment in text data, such as tweets, product reviews, and support tickets, you can understand how customers feel about your brand, product, or service, and gain insights that lead to data-driven decisions.

Sentiment Analysis deals with analyzing emotions and the perspective of a speaker or an author from a given piece of text. "Sentiment analysis or opinion mining refers to the appliance of language process, linguistics, and text analytics to spot and extract subjective information in supply materials". This field of technology deals with analyzing and predicting the hidden information keep within the text. This hidden information gives valuable insights regarding user's intentions, style and odds. Sentiment Analysis specializes in categorizing the text at the extent of subjective and objective nature. Judgement indicates that the text bears opinion content where's perspicacity indicates that the text is while not opinion content

Some examples-

1. Subjective- This motion picture by tom cruise and Angelina jolie is great. (This sentence has an opinion; it talks regarding the motion picture and also the writer's emotions regarding same "great" and thence its subjective

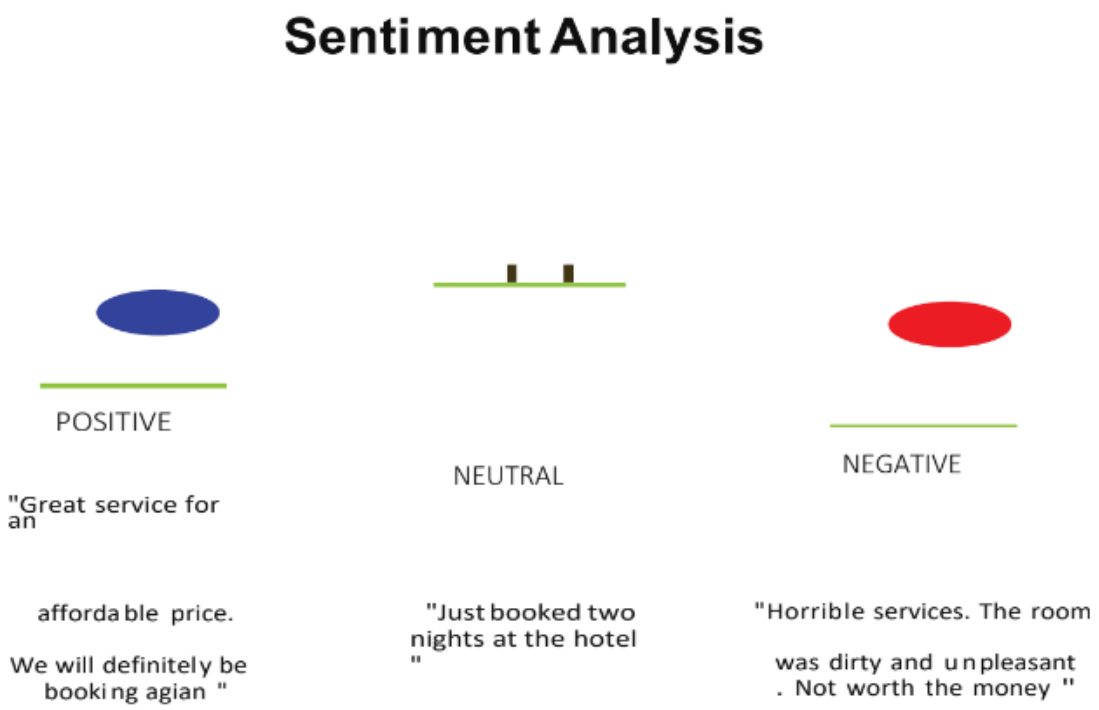
2. Objective- This motion picture stars tom cruise and Angelina.

(This sentence may be a reality, general information instead of an opinion or a read of some individual and thence its objective)

The subjective text may be additional categorized into three broad classes supported the emotions expressed within the text.

1. Positive- I like to look at Star series.
2. Negative- The movie was awful.
3. Neutral- I typically get hungry by evening. (This sentence has users views, emotions hence it's subjective however because it doesn't have any positive or negative polarity therefore it's neutral).

Example:



Method 1: Using Positive and Negative Word Count-With Normalization for Calculating Sentiment Score.

In this method, we will calculate the Sentiment Scores by classifying and counting the Negative and Positive words from the given text and taking the ratio of the difference of Positive and Negative Word Counts and Total Word Count. We will be using the Amazon Cell Phone Reviews dataset from Kaggle.

Code and Output:

```
import pandas as pd

import seaborn as sns

import matplotlib.pyplot as plt

from nltk.sentiment.vader import SentimentIntensityAnalyzer

sentiments = SentimentIntensityAnalyzer()

data = pd.read_csv("/kaggle/input/amazon-cell-phones-reviews/20191226-reviews.csv")

print(data.head())
```

```
         asin           name  rating              date  verified  \
0  B0000SX2UC          Janet       3   October 11, 2005     False
1  B0000SX2UC     Luke Wyatt       1    January 7, 2004     False
2  B0000SX2UC         Brooke       5  December 30, 2003     False
3  B0000SX2UC  amy m. teague       3     March 18, 2004     False
4  B0000SX2UC  tristazbimmer       4    August 28, 2005     False

                                  title  \
0                 Def not best, but not worst
1                 Text Messaging Doesn't Work
2                           Love This Phone
3                    Love the Phone, BUT...!
4  Great phone service and options, lousy case!

                                          body  helpfulVotes
0  I had the Samsung A600 for awhile which is abs...          1.0
1  Due to a software issue between Nokia and Spri...         17.0
2  This is a great, reliable phone. I also purcha...          5.0
3  I love the phone and all, because I really did...          1.0
4  The phone has been great for every purpose it ...          1.0
```

print(data.describe())

	rating	helpfulVotes
count	67986.000000	27215.000000
mean	3.807916	8.229690
std	1.582906	31.954877
min	1.000000	1.000000
25%	3.000000	1.000000
50%	5.000000	2.000000
75%	5.000000	5.000000
max	5.000000	990.000000

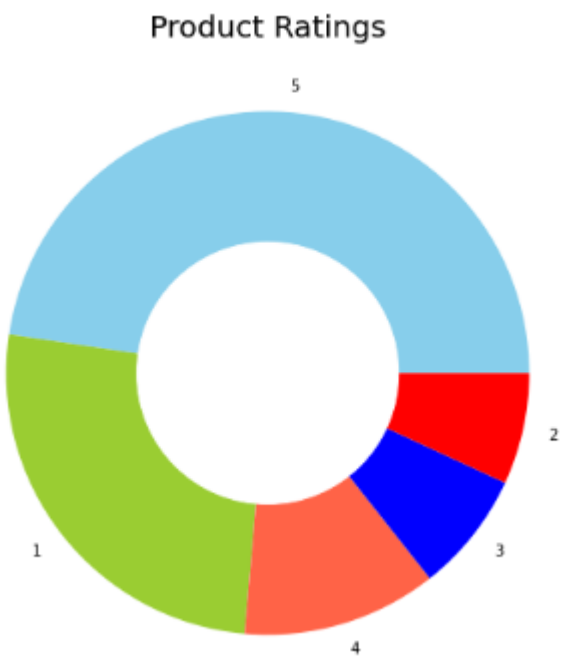
data = data.dropna()

data

	asin	name	rating	date	verified	title	body	helpfulVotes
0	B0000SX2UC	Janet	3	October 11, 2005	False	Def not best, but not worst	I had the Samsung A600 for awhile which is abs...	1.0
1	B0000SX2UC	Luke Wyatt	1	January 7, 2004	False	Text Messaging Doesn't Work	Due to a software issue between Nokia and Spri...	17.0
2	B0000SX2UC	Brooke	5	December 30, 2003	False	Love This Phone	This is a great, reliable phone. I also purcha...	5.0
3	B0000SX2UC	amy m. teague	3	March 18, 2004	False	Love the Phone, BUT...!	I love the phone and all, because I really did...	1.0
4	B0000SX2UC	tristazbimmer	4	August 28, 2005	False	Great phone service and options, lousy case!	The phone has been great for every purpose it ...	1.0
...	...	...	...	...	...	...	...	...
67978	B081H6STQQ	Rock Edge	5	July 16, 2019	False	Candy bar phone is back!	Update 8/14/19 5 stars now! I've been using th...	12.0
67979	B081H6STQQ	Cindy Cowles	3	July 17, 2019	False	Updated review	Update: If you like sending and receiving pict...	2.0
67980	B081H6STQQ	Los Kositos	5	October 7, 2019	False	From iPhone to Android and loving it	I love my new phone. I've been dying to do a r...	1.0
67981	B081H6STQQ	jande	5	August 16, 2019	False	Awesome Phone, but finger scanner is a big mis...	I love the camera on this phone. The screen is...	1.0
67982	B081H6STQQ	2cool4u	5	September 14, 2019	False	Simply Amazing!	I've been an Xperia user for several years and...	1.0

27206 rows × 8 columns

```
ratings = data["rating"].value_counts()
numbers = ratings.index
quantity = ratings.values
custom_colors = ["skyblue", "yellowgreen", 'tomato', "blue", "red"]
plt.figure(figsize=(10, 8))
plt.pie(quantity, labels=numbers, colors=custom_colors)
central_circle = plt.Circle((0, 0), 0.5, color='white')
fig = plt.gcf()
fig.gca().add_artist(central_circle)
plt.rc('font', size=12)
plt.title("Product Ratings", fontsize=20)
plt.show()
```




```
sentiments = SentimentIntensityAnalyzer()
data["Positive"] = [sentiments.polarity_scores(i)["pos"] for i in data["body"]]
data["Negative"] = [sentiments.polarity_scores(i)["neg"] for i in data["body"]]
data["Neutral"] = [sentiments.polarity_scores(i)["neu"] for i in data["body"]]
print(data.head())
```

	asin	name	rating	date	verified	title	body	helpfulVotes	Positive	Negative	Neutral
0	B0000SX2UC	Janet	3	October 11, 2005	False	Def not best, but not worst	I had the Samsung A600 for awhile which is abs...	1.0	0.105	0.080	0.816
1	B0000SX2UC	Luke Wyatt	1	January 7, 2004	False	Text Messaging Doesn't Work	Due to a software issue between Nokia and Spri...	17.0	0.104	0.020	0.876
2	B0000SX2UC	Brooke	5	December 30, 2003	False	Love This Phone	This is a great, reliable phone. I also purcha...	5.0	0.103	0.051	0.846
3	B0000SX2UC	amy m. teague	3	March 18, 2004	False	Love the Phone, BUT...!	I love the phone and all, because I really did...	1.0	0.156	0.000	0.844
4	B0000SX2UC	tristazbimmer	4	August 28, 2005	False	Great phone service and options, lousy case!	The phone has been great for every purpose it ...	1.0	0.121	0.066	0.814

```
x = sum(data["Positive"])
y = sum(data["Negative"])
z = sum(data["Neutral"])


def sentiment_score(a, b, c):
    if (a>b) and (a>c):
        print("Positive 😊 ")
    elif (b>a) and (b>c):
        print("Negative 😠 ")
    else:
        print("Neutral 😐 ")
sentiment_score(x, y, z)
```

```
Neutral 😐
```

```
print("Positive: ", x)
print("Negative: ", y)
print("Neutral: ", z)
```

```
Positive:  5022.301000000003
Negative:  1756.803999999994
Neutral:  20417.808000000085
```

Method 2: Using Positive and Negative Word Counts - With Semi Normalization to calculate Sentiment Score

In this method, we calculate the sentiment score by evaluating the ratio of Count of Positive Words and Count of Negative Words + 1. Since there is no difference of values involved , the sentiment value will always be more than 0. Also, adding 1 in the denominator would save from Zero Division Error. Let's start with the implementation. The implementation code will remain the same till the Negative and Positive Word Count from the previous part with a difference that this time we don't need the total word count, thus will be omitting that part.

Code:

```
def calculate_sentiment(text):
    positive_words = ["good", "great", "excellent", "awesome", "love"]
    negative_words = ["bad", "poor", "terrible", "awful", "hate"]

    # Tokenize the text into words
    words = text.lower().split()

    # Count positive and negative words
    positive_count = sum(word in positive_words for word in words)
    negative_count = sum(word in negative_words for word in words)

    # Calculate sentiment score (semi-normalized ratio of positive to negative word counts)
    sentiment_score = positive_count / (negative_count + 1)
```

```
    return sentiment_score

# Example usage:
text = "The phone is good and has excellent features, but the battery life is poor."
sentiment_score = calculate_sentiment(text)
print("Sentiment Score:", sentiment_score)
```

Output:

```
Sentiment Score: 2.0
```

Conclusion-

Techniques used for sentimental analysis are

Various techniques are utilized for sentiment analysis, encompassing lexicon-based approaches that rely on predefined word dictionaries to machine learning methods such as Support Vector Machines (SVM), Naive Bayes, and Neural Networks, which train classifiers on labeled datasets. Additionally, deep learning techniques like Recurrent Neural Networks (RNNs) and Convolutional Neural Networks (CNNs) are employed to capture complex patterns in text data. Aspect-based sentiment analysis focuses on identifying sentiment associated with specific attributes mentioned in the text, while hybrid approaches combine multiple techniques to enhance accuracy and robustness. These methods are applied across diverse domains, including social media, product reviews, customer feedback, and news articles, to extract valuable insights about people's opinions, emotions, and attitudes.