# SENTIMENT ANALYSIS FOR MARKETING

## PHASE-5

## Project Documentation & Submission

BALACHANDRAN

## Introduction:

Sentiment analysis, a crucial component of modern marketing strategies, is the process of evaluating and interpreting the emotional tone and attitude expressed in text data, often derived from customer feedback, reviews, social media posts, and other sources. This technique provides businesses with valuable insights into how their products, services, or brand are perceived by the public.

In the context of marketing, sentiment analysis helps companies understand consumer sentiment and preferences, enabling them to make data-driven decisions and optimize their strategies. By automatically categorizing opinions as positive, negative, or neutral, sentiment analysis empowers marketers to:

**Monitor Brand Reputation:** Track and assess online mentions, reviews, and social media conversations to gauge how the public perceives the brand.

**Customer Feedback Analysis:** Analyze customer reviews and feedback to identify areas for improvement and address issues promptly.

**Competitive Intelligence:** Compare and benchmark your brand against competitors, gaining insights into strengths and weaknesses.

**Product Development:** Use sentiment analysis to inform product development by understanding customer needs and preferences.

**Campaign Effectiveness:** Evaluate the success of marketing campaigns and adjust strategies based on real-time feedback.

**Personalized Marketing:** Tailor marketing messages and offers to individual preferences and sentiment.

**Crisis Management**: Detect and address potential crises or negative publicity early on.

Sentiment analysis relies on natural language processing and machine learning algorithms to automate the sentiment classification process. It can be performed using specialized software, APIs, or custom-built models, depending on the needs and resources of the marketing team.

sentiment analysis in marketing is a powerful tool that harnesses the wealth of unstructured text data available online to provide actionable insights for brand management, customer engagement, and decision-making. It allows marketers to stay attuned to consumer sentiment, adapt strategies, and ultimately drive business growth.

## Tools and Software used in Sentiment Analysis For Marketing:

Sentiment analysis in marketing can be accomplished using various tools and software. Here are some popular options:

1. **Python Libraries:**

-NLTK (Natural Language Toolkit): It provides tools for working with human language data and is often used for text preprocessing.

-TextBlob: A simple library for processing textual data, including sentiment analysis.

2. **Machine Learning Frameworks:**

- Scikit-learn: It offers various machine learning algorithms that can be used for sentiment analysis.

- TensorFlow and PyTorch: These deep learning frameworks are suitable for more advanced sentiment analysis models.

3. **Commercial Sentiment Analysis Tools:**

- IBM Watson Natural Language Understanding: Offers sentiment analysis as part of its suite of NLP tools.

- SAS Sentiment Analysis: Provides advanced sentiment analysis and text mining capabilities.

**4.Social Media Listening Tools:**

- Tools like Brandwatch, Hootsuite Insights, and Meltwater can monitor social media for brand mentions and sentiment.

**5.Text Analytics Platforms:**

- RapidMiner, Lexalytics, and Aylien offer sentiment analysis as part of their text analytics solutions.

6. **APIs:**

- Services like Google Cloud Natural Language API and Microsoft Azure Text Analytics provide sentiment analysis APIs.

7.**Custom Development:**

- Building custom sentiment analysis models using programming languages like **Python** and libraries like **spaCy** or custom machine learning models for specific business needs.

8.**Dashboard and Reporting Tools:**

- Tools like **Tableau** or **Power BI** can be used to visualize and report on sentiment analysis results.

9.**Data Labeling Tools:**

- Tools like **Prodigy** or **Labelbox** can help in labeling data for training sentiment analysis models.

## Loading data:

IN[]:

```
Import pandas as pd

Df = pd.read_csv("Tweeter.csv")
```

Load the dataset..,

Print(df.head())

OR

o/p[]:

| tweet_id | airline_sen... | airline_sen... | negativere... | negativere... | airline |
|---|---|---|---|---|---|
| 570306133677760513 | neutral | 1.0 | | | Virgin America |
| 570301130888122368 | positive | 0.3486 | | 0.0 | Virgin America |
| 570301083672813571 | neutral | 0.6837 | | | Virgin America |
| 570301031407624196 | negative | 1.0 | Bad Flight | 0.7033 | Virgin America |
| 570300817074462722 | negative | 1.0 | Can't Tell | 1.0 | Virgin America |
| 570300767074181121 | negative | 1.0 | Can't Tell | 0.6842 | Virgin America |
| 570300616901320704 | positive | 0.6745 | | 0.0 | Virgin America |
| 570300248553349120 | neutral | 0.634 | | | Virgin America |
| 570299953286942721 | positive | 0.6559 | | | Virgin America |
| 570295459631263746 | positive | 1.0 | | | Virgin America |
| 570294189143031808 | neutral | 0.6769 | | 0.0 | Virgin America |
| 570289724453216256 | positive | 1.0 | | | Virgin America |
| 570289584061480960 | positive | 1.0 | | | Virgin America |
| 570287408438120448 | positive | 0.6451 | | | Virgin America |

## Preprocessing dataset:

**1.Text Cleaning (or)preprocessing:**

Remove special characters, URLs, and other unwanted elements from the text.

IN[]:

```
Import re
Def clean_text(text):
    Text = re.sub(r'http¥S+', '', text)  # Remove URLs
    Text = re.sub(r'[^A-Za-z0-9]+', ' ', text)  # Remove special characters
    Text = text.lower()  # Convert to lowercase
    Return   text
Clean the dataset..,
Cleaned_text = clean_text(text)
```

o/p[]:

## 2.Tokenization:

Now we will tokenize all the cleaned tweets in our dataset. Tokens are individual terms or words, and tokenization is the process of splitting a string of text into tokens

IN[]:

```
From nltk.tokenize import word_tokenize
```

Def tokenize_text(text):

  Tokens = word_tokenize(text)

  Return tokens

Splits the datasets,..

tokenized_tweet = combi['tidy_tweet'].apply(lambda x: x.split())
tokenized_tweet.head()

Tokens = tokenize_text(cleaned_text)

## 3.**Stop Word Removal:**

Remove common stop words like”and”,”the”,”is”, that do not provide significant information.

IN[]:

From nltk.corpus import stopwords

Def remove_stopwords(tokens):

Stop_words = set(stopwords.words(‘english’))

Filtered_tokens = [word for word in tokens if word.lower() not in stop_words]

Return filtered_tokens

Filtered_tokens = remove_stopwords(tokens)

## 4.**Lemmatization:**

Reduce words to their base form (lemmas).Remove prefixs and suffixs. IN[]:

```
Import spacy

 Nlp = spacy.load(“en_core_web_sm”)

Def lemmatize_text(text):

    Doc = nlp(text)

 Lemmatized_text = ‘ ‘.join([token.lemma_ for token in doc])
```

Return lemmatized_text

Lemmatized_text = lemmatize_text(“ “.join(filtered_tokens))

In [2]:

train =

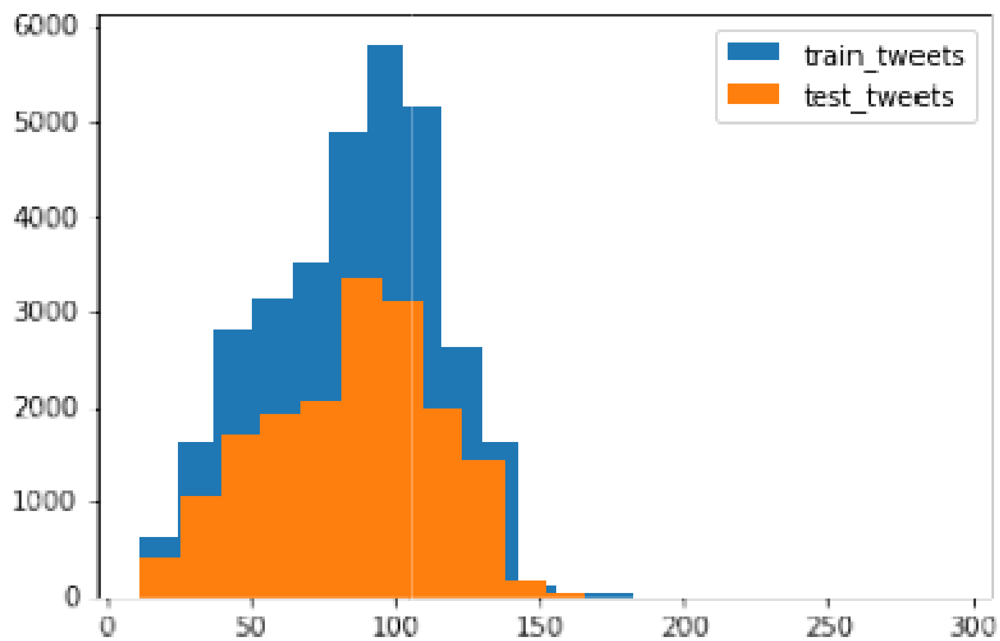
pd.read_csv('train_E6oV3lV.csv') test

=

pd.read_csv('test_tweets_airlines.csv

)

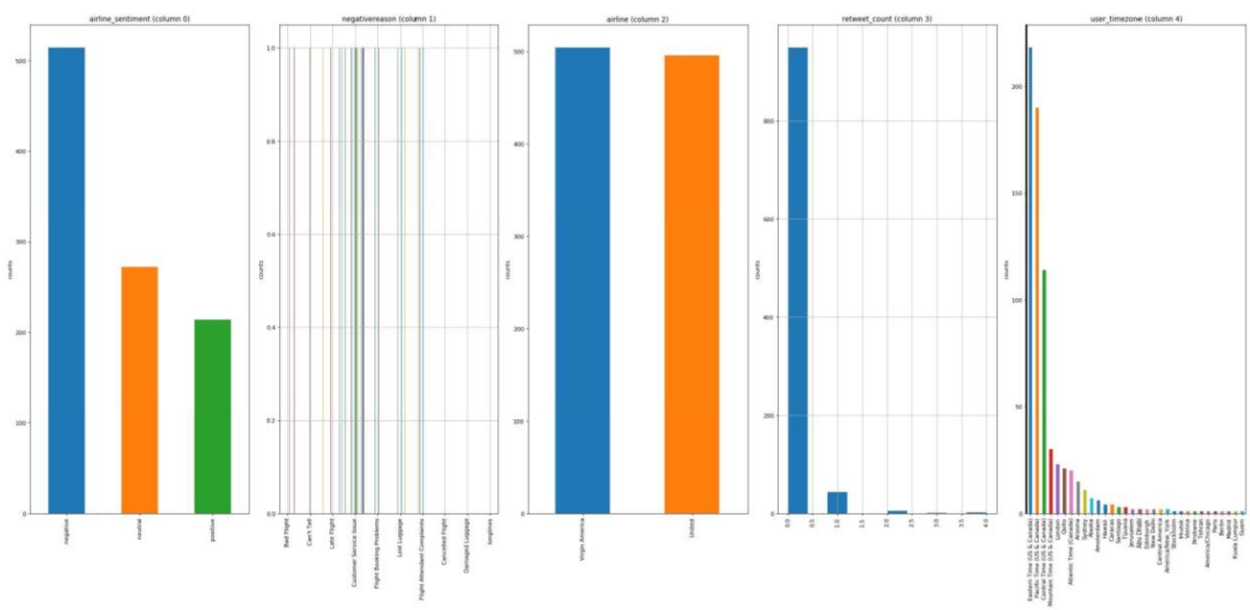
o/p[]:



## Data Inspection :

Let's check out a few tweets.In [3]:

```
train[train['label'] == 0].head(10)
```




## Ensemble Methods:

Bagging: Techniques like Random Forest or Bagged Decision Trees can be employed to create an ensemble of sentiment classifiers. Each classifier is trained on a subset of the data, and their predictions are combined to produce a final sentiment score.

Boosting: Algorithms like AdaBoost or Gradient Boosting can improve sentiment analysis by giving more weight to misclassified data points in each iteration, leading to better overall accuracy.

## Deep Learning Architectures:

**Convolutional Neural Networks (CNNs):** CNNs can be used for sentiment analysis by treating text as an image, converting words into vectors, and using convolutional layers to detect important features.

**Recurrent Neural Networks (RNNs):** RNNs, particularly LSTM (Long Short-Term Memory) and GRU (Gated Recurrent Unit) variants, are effective for sequence modeling, making them suitable for sentiment analysis tasks where the order of words matters.

**Transformer Models:** State-of-the-art models like BERT, GPT, and RoBERTa have revolutionized natural language understanding tasks, including sentiment analysis. They can capture context and nuances in text effectively.

## Bagging:

I/N[]:

```
Accuracy = accuracy_score(y_test, y_pred)

Report = classification_report(y_test, y_pred)

Print(f”Accuracy: {accuracy}”)
```

Print(“Classification Report:\n”, report)

## Boosting:

I/N[]:

Y_pred = ada_boost_classifier.predict(X_test_tfidf)

Accuracy = accuracy_score(y_test, y_pred)

Print(f’Accuracy: {accuracy:.2f}’)

## Recurrent neural network:

I/N[]:

Model.compile(optimizer=’adam’, loss=’binary_crossentropy’, metrics=[‘accuracy’])

Model.fit(X_train, y_train, epochs=5, batch_size=64, validation_data=(X_test, y_test))

Loss, accuracy = model.evaluate(X_test, y_test)

Print(f’Loss: {loss}, Accuracy: {accuracy}’)

Print(f”Text: {text}\nSentiment: {‘Positive’ if sentiment > 0.5 else ‘Negative’}”)

## Convolution neural network:

I/N[]:

Texts = [“This is a positive review.”, “Negative sentiment in this one.”, …]

Model = keras.Sequential

Test_loss, test_acc = model.evaluate(x_test, y_test)

Print(“Test accuracy:”, test_acc)

## BERT:

I/N[]:

Model_name = “bert-base-uncased”  # You can choose different BERT variants

Tokenizer = BertTokenizer.from_pretrained(Tweet)

Model = BertForSequenceClassification.from_pretrained(Tweet)

With torch.no_grad():

Outputs = model(**inputs)

Sentiment_labels = {0: “Negative”, 1: “Neutral”, 2: “Positive”}

Sentiment = sentiment_labels[predicted_label]

Text_to_analyze = “Positive”

Result = analyze_sentiment(text_to_analyze)

Print(f”Sentiment: {result}”)

## RoBERTa:

I/N[]:

Tokenizer = RobertaTokenizer.from_pretrained(Tweet)

Model = RobertaForSequenceClassification.from_pretrained(Tweet)

Logits = outputs.logits

Sentiment_labels = [“Negative”, “Neutral”, “Positive”]

Sentiment = sentiment_labels[predicted_class]

Return sentiment, logits.tolist()

Text = “Positive”

Sentiment, scores = analyze_sentiment(Positive)

Print(f”Sentiment: {sentiment}”)

Print(f”Sentiment Scores: {scores}”)

## GPT-2:

I/N[]:

Model_name = “gpt2” # You can also specify other variants like “gpt2-medium”, “gpt2-large”, etc.

Tokenizer = GPT2Tokenizer.from_pretrained(Tweet)

```
Model = GPT2LMHeadModel.from_pretrained(Tweet)

Input_ids = tokenizer.encode(prompt, return_tensors=”pt”)

Generated_text = tokenizer.decode(output[0],
skip_special_tokens=True)

Print(generated_text)
```

In[]:

```
From mpl_toolkits.mplot3d import Axes3D

From sklearn.preprocessing import StandardScaler

Import matplotlib.pyplot as plt

Import numpy as np

Import os

Import pandas as pd

Print(os.listdir(‘../input’))

nRowsRead =

Df1 = pd.read_csv(‘../input/Tweets.csv’, delimiter=’,’, nrows =
nRowsRead)

Df1.dataframeName = ‘Tweets.csv’

nRow, nCol = df1.shape

print(f’There are {nRow} rows and {nCol} columns’)
```

**O/p[]:**

| | Tweet_id | Airline_sentiment | AS_co |
|---|---|---|---|
| 1 | 5787676523440432 | neutral | 1.000( |
| 2 | 5766756565436587 | positive | 0.784: |

## Advantages:

Sentiment analysis offers several advantages for marketing:

**Customer Insight:** It provides valuable insights into consumer opinions, preferences, and emotions, helping marketers better understand their target audience.

**Real-time Feedback:** Sentiment analysis allows for the immediate detection of trends and changes in public sentiment, enabling rapid responses to emerging issues or opportunities.

**Improved Customer Engagement:** Marketers can use sentiment analysis to personalize interactions with customers based on their sentiments and preferences.

**Competitive Analysis**: It helps assess how your brand compares to competitors, identifying areas for improvement or differentiation.

**Product Development:** Sentiment analysis informs product development by identifying features or improvements that customers desire.

**Campaign Effectiveness:** Marketers can evaluate the success of marketing campaigns and adjust strategies in real-time to maximize their impact.

**Reputation Management:** It aids in monitoring and managing brand reputation by identifying and addressing negative sentiment early.

**Cost Efficiency:** Automation through sentiment analysis reduces the time and resources required for manual analysis of large volumes of text data.

**Targeted Marketing:** It enables more precise targeting of marketing efforts by tailoring messages and offers to different sentiment segments.

**Data-Driven Decision Making:** Marketers can base their strategies on data-driven insights rather than guesswork, leading to more effective campaigns.

**Crisis Management:** Sentiment analysis helps in early detection of potential crises, allowing for proactive measures to mitigate damage.

**Scalability:** It can handle vast amounts of unstructured text data, making it scalable for businesses of all sizes.

Sentiment analysis empowers marketers with a deeper understanding of their customers, the ability to adapt strategies in real-time, and the capacity to make informed decisions that enhance brand reputation and drive business growth.

## Disadvantages:

While sentiment analysis offers numerous benefits for marketing, it also comes with some disadvantages and limitations:

**Ambiguity and Context:** Sentiment analysis struggles with understanding the nuances of language, sarcasm, irony, and context, which can lead to misinterpretations of sentiment.

**Subjectivity:** Sentiment analysis often relies on pre-defined sentiment categories (positive, negative, neutral), which may not capture the full range of opinions and emotions expressed by customers.

**Accuracy and Reliability:** The accuracy of sentiment analysis can vary depending on the quality of data, language, and the complexity of the task. It may produce false positives or false negatives.

**Multilingual Challenges:** Sentiment analysis may not perform equally well in all languages, and accuracy can decrease when analyzing languages with less training data available.

**Domain Specificity:** Pre-trained sentiment analysis models might not be tailored to a specific industry or domain, leading to inaccuracies when applied to specialized subjects.

**Data Preprocessing:** The need for data preprocessing and cleaning can be time-consuming, especially when dealing with unstructured and noisy data sources.

**Lack of Depth:** Sentiment analysis typically categorizes sentiment as positive, negative, or neutral, but it may not capture more complex emotions or provide in-depth insights.

**Human Bias:** Sentiment analysis models can inherit biases present in the training data, potentially leading to unfair or misleading results.

**Limited to Text:** Sentiment analysis is primarily designed for text data and does not directly analyze visual or auditory content, limiting its scope.

**Constantly Changing Language:** The evolution of language and the emergence of new slang or expressions can challenge the effectiveness of sentiment analysis models.

**Integration Complexity:** Implementing sentiment analysis tools and APIs into existing marketing systems and workflows may require technical expertise and effort.

**Data Privacy and Ethics:** Handling customer data for sentiment analysis raises privacy and ethical concerns, requiring careful data management and compliance with regulations.

While sentiment analysis is a valuable tool for marketing, it is not without its limitations and challenges. Marketers should be aware of these drawbacks and use sentiment analysis in conjunction with other research methods to make well-informed decisions.

## Benefits Of Sentiment Analysis For Marketing:

Sentiment analysis can be a valuable tool for marketing in a number of ways, including:

**Understanding customer needs and preferences**: Sentiment analysis can help marketers understand what customers like and dislike about their products, services, and brand. This information can be used to develop new products and services, improve existing ones, and create targeted marketing campaigns.

**Identifying and responding to customer issues:** Sentiment analysis can be used to identify customer complaints and feedback. This information can be used to improve customer service, resolve customer issues, and prevent future problems.

**Improving brand reputation:** Sentiment analysis can be used to track customer sentiment towards a brand over time. This information can be used to identify areas where the brand needs to improve and to measure the effectiveness of marketing campaigns.

**Analyzing competitor performance:** Sentiment analysis can be used to analyze customer sentiment towards competitor brands. This information can be used to identify areas where the brand is performing well and areas where it could improve.

**Developing more effective marketing campaigns:** Sentiment analysis can be used to develop more effective marketing campaigns by tailoring messages to the specific needs and preferences of different customer segments.

Specific examples of how marketers can use sentiment analysis:

A social media marketer can use sentiment analysis to track customer reactions to a new product launch and identify any negative feedback that needs to be addressed.

A customer service representative can use sentiment analysis to identify and prioritize customer complaints.

A product manager can use sentiment analysis to understand what features customers are most interested in and to identify any areas where the product needs to be improved.

A marketing manager can use sentiment analysis to measure the effectiveness of different marketing campaigns and to identify areas where the campaigns can be improved.

Overall, sentiment analysis is a powerful tool that can help marketers better understand their customers, improve their products and services, and develop more effective marketing campaigns.

In addition to the above, sentiment analysis can also be used for marketing research purposes. For example, marketers can use sentiment analysis to:

- Identify trends in customer opinion.
- Understand how customer sentiment varies across different demographics and regions.

- Assess the impact of marketing campaigns on customer sentiment.

Sentiment analysis is a relatively new technology, but it is quickly becoming an essential tool for marketers who want to stay ahead of the curve.

## Conclusion:

Sentiment analysis is a powerful tool for modern marketing, offering a data-driven approach to understanding and harnessing the emotions, opinions, and attitudes of consumers. This technology enables marketers to gain valuable insights, enhance customer experiences, and drive business success. By monitoring and analyzing sentiment in real-time, marketing professionals can adapt strategies, protect brand reputation, and engage with customers in a more personalized and effective manner. Sentiment analysis is a key asset for staying competitive, making informed decisions, and ultimately achieving marketing goals in today’s dynamic and customer-centric landscape. Its benefits extend from customer understanding to crisis management, ensuring that businesses can thrive in the ever-evolving world of marketing.