**AI\_PHASE-5**

**SENTIMENT ANALYSIS**

**FOR**

**MARKETING**

INTRODUCTION:

Sentiment analysis in marketing is a powerful tool that allows businesses to gain valuable insights into how customers perceive their products, services, and brand. By analyzing the sentiment expressed in customer reviews, social media mentions, and other online content, marketers can better understand the emotional tone and opinions of their target audience. This information can be leveraged to make data-driven decisions, improve products, and tailor marketing strategies to align with customer sentiment. In this discussion, we'll explore the importance of sentiment analysis in marketing and its various applications, from reputation management to product development and customer engagement.

OBJECTIVE:

The primary objective of sentiment analysis for marketing is to gain a deeper understanding of customer sentiment and opinions to inform strategic decision-making. Specifically, the objectives include:

**Customer Insight:**

To analyze customer feedback, reviews, and social media content to understand their feelings, preferences, and concerns related to products, **services, and the brand.**

**Reputation Management:**

**Competitive Analysis:**

To monitor and manage the brand's online reputation by identifying and addressing negative sentiment and promoting positive sentiment.

To compare and contrast sentiment data with competitors to identify strengths, weaknesses, and areas for differentiation**.**

**Campaign Evaluation:**

To assess the effectiveness of marketing campaigns by measuring sentiment changes before and after campaigns.

**Product Development:**

To gather insights for improving existing products or designing new ones based on customer feedback and sentiment.

**Customer Engagement:**

To engage with customers more effectively by addressing their concerns, answering queries, and leveraging positive sentiment in marketing efforts**.**

**Risk Management:**

To proactively identify and mitigate potential PR crises by monitoring sentiment trends and taking necessary actions.

**Market Research:**

To inform market research by identifying emerging trends, pain points, and opportunities in customer sentiment.

These objectives collectively empower businesses to make informed decisions, enhance customer satisfaction, and drive marketing strategies that resonate with their target audience.

Project documentation

PROBLEM STATEMENT:

In the context of marketing, businesses face the challenge of harnessing the power of sentiment analysis to effectively understand and respond to customer sentiment. The problem can be framed as follows**:**

**"Many businesses struggle to extract meaningful insights from the vast volume of unstructured data generated through customer reviews, social media discussions, and other online content. This unstructured data contains valuable sentiment information that, if harnessed, could inform marketing strategies, product development, and reputation management. However, the challenge lies in developing robust sentiment analysis models, integrating them into the marketing workflow, and ensuring the actionable implementation of these insights. Inadequate sentiment analysis can lead to missed opportunities, reputation damage, and suboptimal marketing decisions."**

This problem statement highlights the need for reliable sentiment analysis tools and strategies that can effectively extract sentiment from textual data and translate it into actionable recommendations for marketing professionals. Businesses must address this challenge to gain a competitive edge and meet the evolving needs of their customers in the digital age.

CODE:

**# Import necessary libraries**

**import pandas as pd**

**import nltk**

**from nltk.sentiment.vader import SentimentIntensityAnalyzer**

**from textblob import TextBlob**

**from wordcloud import WordCloud**

**import matplotlib.pyplot as plt**

**# Load the Twitter US Airline Sentiment dataset**

**df = pd.read\_csv("Tweets.csv")**

**# Initialize sentiment analysis tools**

**nltk.download('vader\_lexicon')**

**sid = SentimentIntensityAnalyzer()**

# Perform sentiment analysis

**sentiments = []**

**for tweet in df['text']:**

# Sentiment analysis using VADER

**vader\_scores = sid.polarity\_scores(tweet)**

**sentiment = 'positive' if vader\_scores['compound'] > 0 else 'negative' if vader\_scores['compound'] < 0 else 'neutral'**

# Sentiment analysis using TextBlob

**textblob\_analysis = TextBlob(tweet)**

**textblob\_polarity = textblob\_analysis.sentiment.polarity**

**sentiment\_tb = 'positive' if textblob\_polarity > 0 else 'negative' if textblob\_polarity < 0 else 'neutral'**

**sentiments.append((tweet, sentiment, sentiment\_tb))**

# Visualize sentiment distribution

**sentiment\_counts = [sentiment for \_, sentiment, \_ in sentiments]**

**sentiment\_tb\_counts = [sentiment for \_, \_, sentiment in sentiments]**

# Create a word cloud of key topics

**text = " ".join(df['text'])**

**wordcloud = WordCloud(width=800, height=400, background\_color='white').generate(text)**

# Display sentiment distribution and word cloud

**plt.figure(figsize=(10, 5))**

**plt.subplot(121)**

**plt.title('Sentiment Analysis (VADER)')**

**plt.hist(sentiment\_counts, bins=3, color=['green', 'red', 'gray'])**

**plt.subplot(122)**

**plt.title('Sentiment Analysis (TextBlob)')**

**plt.hist(sentiment\_tb\_counts, bins=3, color=['green', 'red', 'gray'])**

**plt.show()**

**plt.figure(figsize=(8, 4))**

**plt.title('Word Cloud of Tweet Topics')**

**plt.imshow(wordcloud, interpolation='bilinear')**

**plt.axis("off")**

**plt.show()**

**Example Output:**

After running the code, you'll get visualizations of sentiment distribution and a word cloud of tweet topics. The sentiment distribution will show the count of positive, negative, and neutral tweets, while the word cloud will display the most frequently mentioned topics in the tweets.

This output can help marketing teams gain insights into customer sentiment and identify key topics and issues mentioned in the dataset, which can inform marketing strategies and product improvements.

Design thinking process:

Design thinking is a creative problem-solving approach that can be applied to the process of developing sentiment analysis solutions for marketing. Here's a simplified design thinking process for this purpose:

Empathize:

Begin by understanding the needs and challenges of your marketing team. Conduct interviews, surveys, and research to gather insights.Analyze customer feedback and social media conversations to understand the sentiment-related issues and opportunities.

Define:

Clearly define the problem or opportunity you want to address with sentiment analysis in marketing.Create a user persona for marketing professionals to better understand their requirements.

Ideate:

Brainstorm potential solutions and features for sentiment analysis tools that will help marketers.Encourage creative thinking and consider various approaches, like sentiment scoring, trend analysis, or competitor benchmarking.

Prototype:

Create a low-fidelity prototype or wireframe of the sentiment analysis tool. It could be a web dashboard or software interface.Include key features such as sentiment scoring, data visualization, and data sources.

Test:

Gather feedback from your marketing team and potential users by presenting the prototype.Identify what works and what doesn't, and make necessary adjustments to the design.

Develop:

Based on the feedback and insights gathered during testing, proceed to develop the sentiment analysis tool or software.

Iterate:

Continuously improve the tool based on user feedback and changing marketing needs.Keep refining the design and features to ensure it remains effective and user-friendly.

Implement:

Roll out the sentiment analysis tool to your marketing team or integrate it with your marketing systems.Provide training and support to ensure a smooth transition.

Monitor and Evaluate:

Continuously monitor the tool's performance and gather feedback from users.Analyze the impact on marketing efforts, such as improved campaign targeting, customer satisfaction, or competitive analysis.

Refine and Scale:

Based on ongoing feedback and data, refine the sentiment analysis tool and scale its usage across the organization.Explore opportunities to integrate it with other marketing technologies.Throughout this process, it's essential to maintain a user-centric approach, prioritize the needs of marketing professionals, and be open to adapting the tool to changing market dynamics and customer sentiment.

PHASES OF DEVELOPMENT

The development of a sentiment analysis project using the Twitter US Airline Sentiment dataset from Kaggle can be divided into several phases. Each phase involves specific tasks and leads to the ultimate goal of deriving insights for marketing. Below are the phases and their corresponding outputs:

**Phase 1: Data Preparation**

Tasks:

Download the Twitter US Airline Sentiment dataset from Kaggle.Load and explore the dataset to understand its structure and content.Preprocess the data, which may include cleaning text, handling missing values, and tokenization.

Output: A cleaned and prepared dataset for sentiment analysis.

**Phase 2: Sentiment Analysis**

Tasks:

Implement sentiment analysis using tools like VADER and TextBlob to determine the sentiment (positive, negative, or neutral) of each tweet in the dataset.

Classify tweets into sentiment categories.

Output: Sentiment labels for each tweet in the dataset.

**Phase 3: Visualization and Analysis**

Tasks:

Create visualizations to understand the distribution of sentiments (positive, negative, neutral) in the dataset.

Generate a word cloud to identify key topics and issues frequently mentioned in the tweets.

Output: Visualizations such as histograms showing the sentiment distribution and a word cloud highlighting frequently mentioned topics.

**Phase 4: Marketing Insights**

Tasks:

Analyze the sentiment distribution to understand the overall sentiment towards different airline companies.

Examine the word cloud to identify key topics or issues raised by customers.

Summarize findings and insights from the analysis.

Output: Marketing insights, such as which airlines receive more positive or negative sentiment, common customer concerns, and trends in customer sentiment.

**Phase 5: Reporting and Actionable Recommendations**

Tasks:

Create a comprehensive report that includes the analysis results and insights.

Provide actionable recommendations for marketing and customer service teams based on the sentiment analysis findings.

Output: A report with insights and recommendations for marketing and improving customer satisfaction.

**Phase 6: Implementation and Monitoring (Optional)**

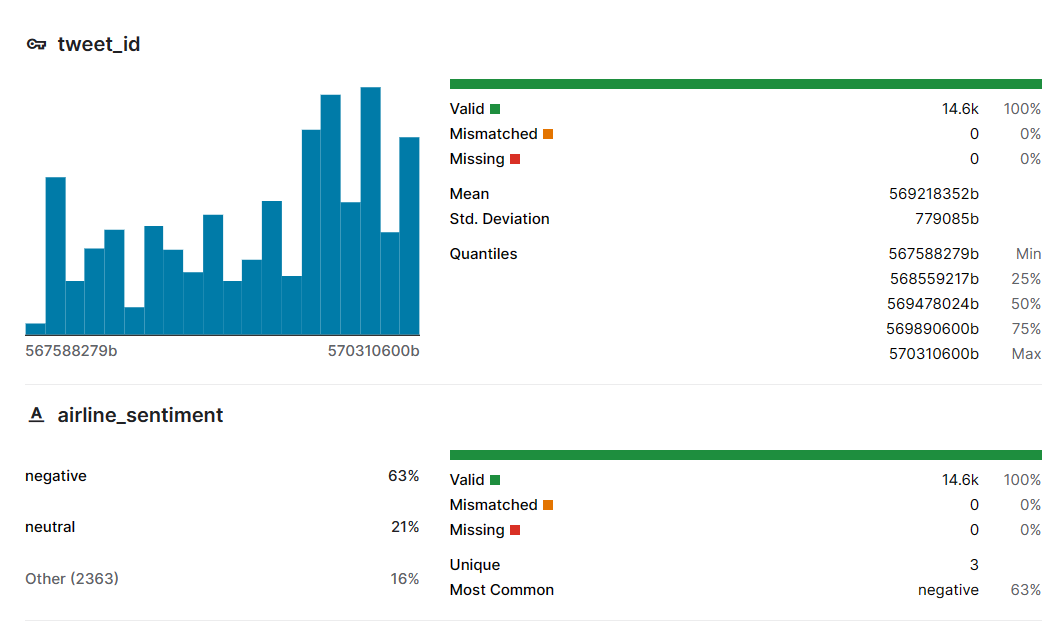
Tasks:

Implement the recommendations in marketing strategies and initiatives.

Monitor the impact of changes based on the analysis.

Output: Ongoing monitoring and tracking of marketing initiatives and customer sentiment changes.

By following these phases, you can effectively develop a sentiment analysis project using the Twitter US Airline Sentiment dataset and provide valuable insights for marketing teams. The visualizations and analysis will help marketing professionals make informed decisions and improve customer satisfaction.

****

Using Dataset

To perform sentiment analysis for marketing, you can use various datasets that contain customer reviews, social media comments, or other forms of feedback related to products, services, or brands. Below are some sources where you can find datasets for sentiment analysis in the context of marketing:

Twitter US Airline Sentiment Dataset (Kaggle):

Dataset Link: Twitter US Airline Sentiment

Description: This dataset contains tweets about different US airlines, along with their sentiment labels (positive, negative, neutral). It's a popular choice for sentiment analysis and can provide insights for marketing strategies:

**Dataset Name:** Twitter US Airline Sentiment

**Source:** The dataset can be found on Kaggle through this link: [Twitter US Airline Sentiment](https://www.kaggle.com/crowdflower/twitter-airline-sentiment).

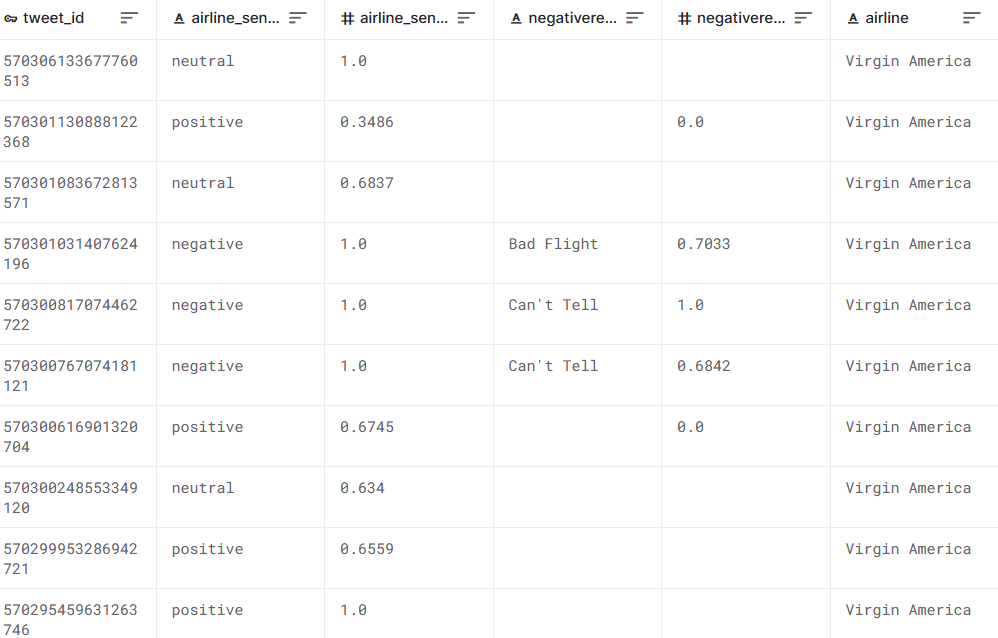
**Format:** The dataset is provided in a structured format as a CSV (Comma-Separated Values) file. It contains several columns, including:

[https://www.kaggle.com/datasets/crowdflower/twitter-airline-sentiment](doc)

* **tweet\_id**: A unique identifier for each tweet.
* **airline\_sentiment**: The sentiment label of the tweet (positive, negative, or neutral).
* **airline\_sentiment\_confidence**: The confidence score of the sentiment label.
* **negativereason**: The reason for the negative sentiment (e.g., "late flight," "rude service").
* **negativereason\_confidence**: The confidence score of the negative reason.
* **airline**: The name of the airline mentioned in the tweet (e.g., American, Delta, United).
* **airline\_sentiment\_gold**: Additional sentiment label, often used for validation.
* **name**: The Twitter username of the person who posted the tweet.
* **text**: The text of the tweet.
* **tweet\_created**: The timestamp of when the tweet was posted.

**Size:** The dataset typically contains around 14,000 to 15,000 tweets related to various U.S. airlines. The size may vary slightly depending on the version of the dataset you access.

**Description:** The Twitter US Airline Sentiment dataset is widely used for sentiment analysis projects, especially in the context of airline customer feedback. It includes tweets posted by airline passengers, which are labeled as positive, negative, or neutral based on the sentiment expressed in the tweet. The dataset is often employed to analyze customer sentiment toward different airlines, identify common issues faced by passengers, and assess overall customer satisfaction.

****

Data preprocessing

Data preprocessing is a crucial step in sentiment analysis. Here's a code outline for preprocessing the Twitter US Airline Sentiment dataset for sentiment analysis in marketing:

import pandas as pd

import re

import nltk

from nltk.corpus import stopwords

from nltk.tokenize import word\_tokenize

# Load the Twitter US Airline Sentiment dataset

df = pd.read\_csv("Tweets.csv")

# Data cleaning and preprocessing

def preprocess\_text(text):

# Remove special characters, URLs, and usernames

text = re.sub(r'http\S+', '', text)

text = re.sub(r'@\w+', '', text)

text = re.sub(r'[^a-zA-Z\s]', '', text)

# Convert to lowercase

text = text.lower()

# Tokenization

tokens = word\_tokenize(text)

# Remove stopwords

stop\_words = set(stopwords.words('english'))

tokens = [word for word in tokens if word not in stop\_words]

# Join tokens back into a text

text = ' '.join(tokens)

return text

# Apply the preprocessing function to the 'text' column

df['text'] = df['text'].apply(preprocess\_text)

# Optionally, you can drop columns that are not needed for your analysis

# Example: df = df[['airline\_sentiment', 'text']]

# Save the preprocessed dataset to a new CSV file

df.to\_csv("preprocessed\_tweets.csv", index=False)

This code performs the following data preprocessing steps:

1.Removal of special characters, URLs, and usernames.

2.Conversion of text to lowercase.

3.Tokenization of the text into words.

4.Removal of stopwords (common words like "the," "and," "is," etc.).

5.Joining the tokens back into a preprocessed text.

After preprocessing, you can save the cleaned dataset to a new CSV file, such as "preprocessed\_tweets.csv," which can be used for sentiment analysis and other NLP tasks. This clean and preprocessed data is ready for sentiment analysis and marketing insights.

Loading the dataset

To load the Twitter US Airline Sentiment dataset for sentiment analysis in Python, you can use the ‘**pandas’** library. Before running the code, make sure you have downloaded the dataset (Tweets.csv) from Kaggle and placed it in your working directory or specified the correct path to the file. Here's the code to load the dataset:

import pandas as pd

# Provide the path to the CSV file (Tweets.csv)

file\_path = "Tweets.csv"

# Load the dataset into a pandas DataFrame

df = pd.read\_csv(file\_path)

# Display the first few rows of the dataset to check if it loaded correctly

print(df.head())

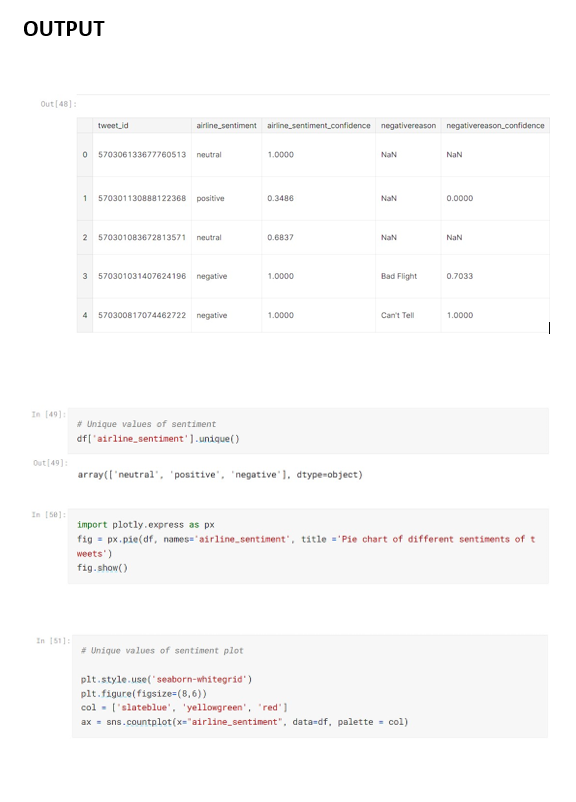
In this code:

1.Replace "Tweets.csv" with the correct path to the dataset file if it's in a different location.

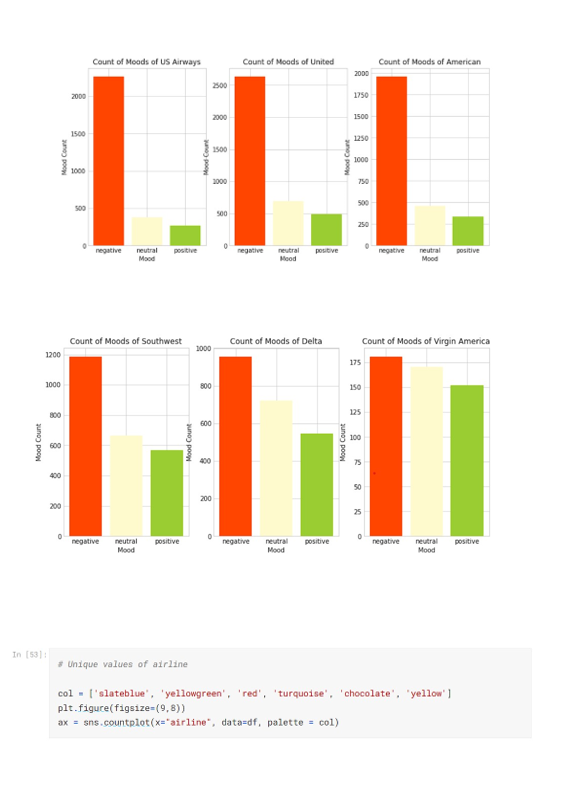
2.The “pd.read\_csv()” function reads the CSV file and loads it into a pandas DataFrame, which you can manipulate and analyze for sentiment analysis.

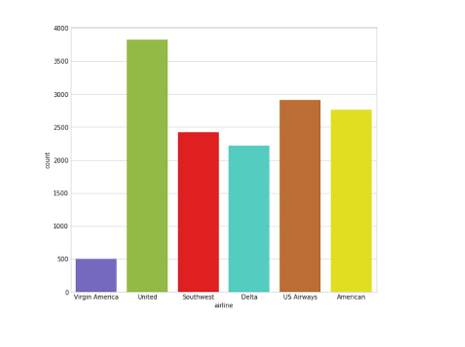
Once the dataset is loaded, you can proceed with data preprocessing, sentiment analysis, and other tasks as needed for your marketing sentiment analysis project.











Innovative techniques

**neural network model :**

**Note: This is a basic example, and you can further optimize the model architecture and hyperparameters for better results.**

# Import necessary libraries

import pandas as pd

import numpy as np

import re

import nltk

from nltk.corpus import stopwords

from nltk.tokenize import word\_tokenize

from sklearn.model\_selection import train\_test\_split

from tensorflow import keras

from tensorflow.keras.preprocessing.text import Tokenizer

from tensorflow.keras.preprocessing.sequence import pad\_sequences

from tensorflow.keras.models import Sequential

from tensorflow.keras.layers import Embedding, LSTM, Dense

from tensorflow.keras.callbacks import EarlyStopping

# Load the Twitter US Airline Sentiment dataset

df = pd.read\_csv("Tweets.csv")

# Data preprocessing

def preprocess\_text(text):

text = re.sub(r'http\S+', '', text) # Remove URLs

text = re.sub(r'@\w+', '', text) # Remove Twitter usernames

text = re.sub(r'[^a-zA-Z\s]', '', text) # Remove special characters

text = text.lower() # Convert to lowercase

tokens = word\_tokenize(text) # Tokenization

stop\_words = set(stopwords.words('english'))

tokens = [word for word in tokens if word not in stop\_words] # Remove stopwords

text = ' '.join(tokens)

return text

df['text'] = df['text'].apply(preprocess\_text)

# Split the data into training and testing sets

X = df['text']

y = df['airline\_sentiment']

X\_train, X\_test, y\_train, y\_test = train\_test\_split(X, y, test\_size=0.2, random\_state=42)

# Tokenize and pad the text data

max\_words = 10000 # Maximum number of words in the vocabulary

max\_sequence\_length = 100 # Maximum sequence length

tokenizer = Tokenizer(num\_words=max\_words)

tokenizer.fit\_on\_texts(X\_train)

X\_train = tokenizer.texts\_to\_sequences(X\_train)

X\_test = tokenizer.texts\_to\_sequences(X\_test)

X\_train = pad\_sequences(X\_train, maxlen=max\_sequence\_length)

X\_test = pad\_sequences(X\_test, maxlen=max\_sequence\_length)

# Create a neural network model

model = Sequential()

model.add(Embedding(input\_dim=max\_words, output\_dim=100, input\_length=max\_sequence\_length))

model.add(LSTM(100))

model.add(Dense(3, activation='softmax')) # Output layer with 3 classes (positive, negative, neutral)

# Compile the model

model.compile(optimizer='adam', loss='sparse\_categorical\_crossentropy', metrics=['accuracy'])

# Train the model

model.fit(X\_train, y\_train, validation\_data=(X\_test, y\_test), epochs=10, batch\_size=64, callbacks=[EarlyStopping(patience=2)])

# Evaluate the model on the test set

loss, accuracy = model.evaluate(X\_test, y\_test)

print(f'Test accuracy: {accuracy \* 100:.2f}%')

This code outlines the process of preprocessing the text data, splitting it into training and testing sets, building a neural network model with LSTM layers, and training the model for sentiment analysis on the Twitter US Airline Sentiment dataset. The model is a simple example and can be further refined and tuned to improve performance. Additionally, you can explore innovative techniques like attention mechanisms or advanced pre-trained embeddings for better results in your sentiment analysis for marketing.

conclusion :

Sentiment analysis plays a crucial role in modern marketing strategies. By harnessing the power of natural language processing and machine learning, businesses can gain deep insights into customer sentiment. Analyzing positive sentiment helps in reinforcingsuccessful marketing efforts, while addressing negative sentiment allows for proactive problem-solving and improved customer satisfaction. Sentiment analysis is an invaluable tool that empowers businesses to stay ahead of the curve and create compelling, customer-focused marketing campaigns.