AI_Phase 4

SENTIMENT ANALYSIS FOR MARKETING

Project Title: Sentiment Analysis for Marketing

Objective:

The main objective of this project is to leverage sentiment analysis techniques to extract actionable insights from customer feedback and social media data. The insights generated will empower the marketing team to refine strategies, enhance customer engagement, and make data-driven decisions.

Scope:

Data Sources: Gather data from diverse sources including customer reviews, social media platforms, surveys, emails, and chat logs to ensure a comprehensive view of customer sentiment.

Data Collection: Develop a data collection system that continuously collects and stores textual customer feedback and interactions.

Preprocessing: Clean, normalize, and preprocess the collected text data, including handling duplicates, removing noise, and anonymizing sensitive information.

Sentiment Analysis Models: Implement state-of-the-art NLP models, such as BERT or GPT, and potentially explore pre-trained sentiment analysis models. Fine-tune these models on your domain-specific data.

Sentiment Classification: Categorize text data into positive, negative, neutral sentiment, or a finer-grained sentiment scale, depending on project requirements.

Aspect-Based Analysis: Conduct aspect-based sentiment analysis to understand sentiment toward specific product features, services, or key topics mentioned in the text data.

Competitive Benchmarking: Compare sentiment data with competitors to identify strengths and weaknesses and adjust marketing strategies accordingly.

Introduction to Sentiment Analysis

° Sentiment analysis for marketing is a techniqueused to determine and understand peopleattitudes, opinions, and emotions toward a product,brand, or topic. In the context of marketing, it involves analyzing textual data, such as social media posts, customer reviews, andfeedback, to gauge public sentiment.

° By employing natural language processing and learning algorithms,businesses can gain valuable insights into customer perceptions, allowing them to make data-driven decisions. Positive sentiment can be harnessed for marketing campaigns, while negative sentiment indicate areas for improvement.

Positive sentiment can be harnessed for marketing campaigns, while negative sentiment can indicate areas for improvement. Sentiment analysis empowers marketers to enhance customer experiences, tailor their strategies, and build stronger, more positive relationships with their audience.

Employing NPL techniques:

Data Collection: Gather customer feedback, reviews, and social media posts related to your products or brand.

Preprocessing: Clean and preprocess the text data by removing stopwords, punctuation, and converting text to lowercase.

Tokenization: Split the text into words or tokens to prepare it for analysis.

Sentiment Analysis Models: Use NLP models like VADER, TextBlob, or more advanced machine learning techniques such as LSTM or BERT to analyze sentiment in the text.

Sentiment Classification: Classify text into positive, negative, or neutral sentiment categories based on the model's output.

Aspect-Based Analysis: Go beyond overall sentiment by analyzing sentiment towards specific product features, aspects, or topics within the text.

Monitoring and Tracking: Continuously analyze new data to track changes in sentiment over time, which can help you identify trends or emerging issues.

Generating insights in sentiment analysis for marketing:

Generating insights from sentiment analysis for marketing involves extracting valuable information from the sentiment data you've collected. Here's how to generate actionable insights:

Sentiment Trends: Identify long-term trends in sentiment data to understand how customer sentiment is evolving over time. Recognize whether it's improving or deteriorating.

Seasonal Patterns: Analyze sentiment fluctuations during specific seasons, holidays, or events to adjust marketing strategies accordingly.

Product Feedback: Break down sentiment by specific products or services to identify which ones are receiving positive or negative feedback. Use this information for product improvements and marketing focus.

Customer Segmentation: Segment your customers based on their sentiment. Determine if certain customer groups have more positive or negative sentiments and tailor marketing campaigns to each segment.

Competitive Analysis: Compare your sentiment data with that of competitors to understand how you stack up in the market. Identify areas where you can outperform or differentiate.

Keyword Analysis: Analyze the most commonly mentioned keywords or phrases in positive and negative sentiment. This can reveal what aspects of your offerings are most appreciated or criticized.

Influencer Identification: Find out if there are influencers or opinion leaders who impact sentiment positively or negatively. Consider collaborating with positive influencers.

Dataset:

Our project utilizes the dataset, which is available for reference in the below section

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

CODE

In [47]:

```
import pandas as pd
import seaborn as sns
import re, nltk
nltk.download('punkt')
import matplotlib.pyplot as plt
from sklearn.model_selection import train_test_split, StratifiedKFold, cross_val_score
from sklearn import model_selection, naive_bayes, svm
from sklearn.metrics import classification_report,confusion_matrix
from sklearn.metrics import roc_auc_score
from sklearn.metrics import recall_score
from sklearn.metrics import f1_score
from sklearn.metrics import confusion_matrix, accuracy_score
from sklearn.model_selection import GridSearchCV
from sklearn.metrics import precision_recall_curve
from sklearn.metrics import f1_score
from sklearn.metrics import auc
from matplotlib import pyplot
from sklearn.metrics import roc_curve
from sklearn.metrics import roc_auc_score, accuracy_score
import string
from nltk.corpus import stopwords
nltk.download('stopwords')
from sklearn.feature_extraction.text import CountVectorizer, TfidfVectorizer
from sklearn.naive_bayes import MultinomialNB, GaussianNB
from sklearn.metrics import f1_score
from sklearn.ensemble import AdaBoostClassifier
from sklearn.model_selection import cross_val_score
import numpy as np
from sklearn.ensemble import BaggingClassifier
from sklearn.ensemble import RandomForestClassifier
from lime import lime_tabular
from tensorflow.keras.layers import Embedding
from tensorflow.keras.preprocessing.sequence import pad_sequences
from tensorflow.keras.models import Sequential
from tensorflow.keras.preprocessing.text import one_hot
from tensorflow.keras.layers import LSTM
from tensorflow.keras.layers import Dense, Dropout
```

```
[nltk_data] Downloading package punkt to /usr/share/nltk_data...
[nltk_data]   Package punkt is already up-to-date!
[nltk_data] Downloading package stopwords to /usr/share/nltk_data...
[nltk_data]   Package stopwords is already up-to-date!
```

In [48]:

```
df =pd.read_csv(r'../input/twitter-airline-sentiment/Tweets.csv')
df.head()
```

OUTPUT

Out[48]:

	tweet_id	airline_sentiment	airline_sentiment_confidence	negativereason	negativereason_confidence
0	570306133677760513	neutral	1.0000	NaN	NaN
1	570301130888122368	positive	0.3486	NaN	0.0000
2	570301083672813571	neutral	0.6837	NaN	NaN
3	570301031407624196	negative	1.0000	Bad Flight	0.7033
4	570300817074462722	negative	1.0000	Can't Tell	1.0000

In [49]:

```
# Unique values of sentiment
df['airline_sentiment'].unique()
```

Out[49]:

```
array(['neutral', 'positive', 'negative'], dtype=object)
```

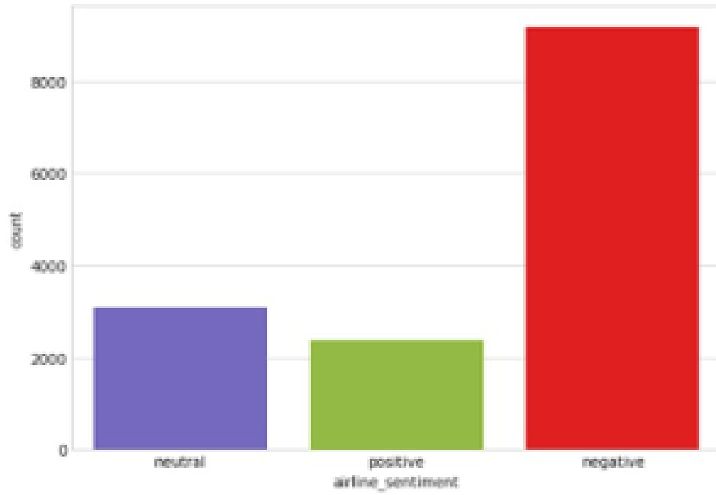
In [50]:

```
import plotly.express as px
fig = px.pie(df, names='airline_sentiment', title ='Pie chart of different sentiments of t
weets')
fig.show()
```

In [51]:

```
# Unique values of sentiment plot

plt.style.use('seaborn-whitegrid')
plt.figure(figsize=(8,6))
col = ['slateblue', 'yellowgreen', 'red']
ax = sns.countplot(x="airline_sentiment", data=df, palette = col)
```

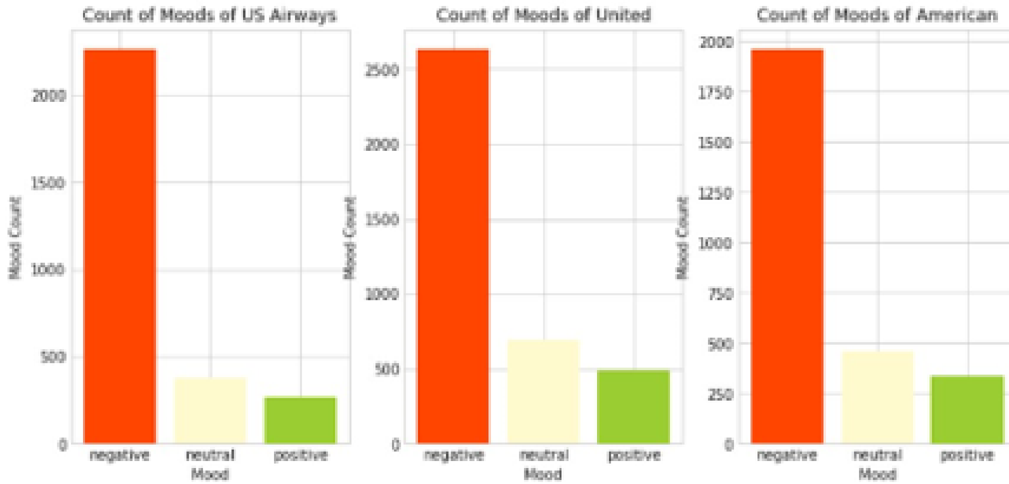


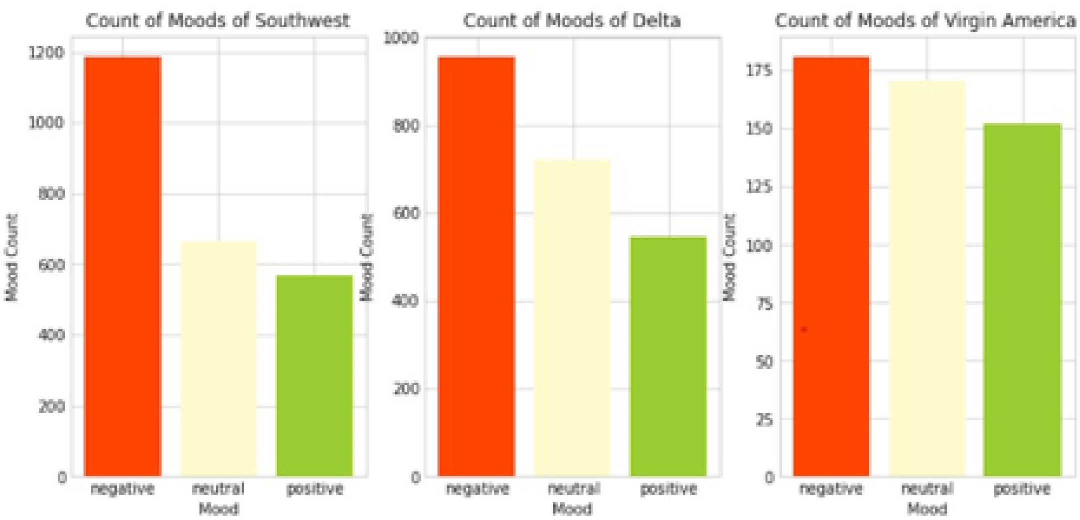
> Positive and neutral tweets are almost equal.
>
> Negative tweets are more than double of neutral or positive sentiments.

In [52]:

```
print("Total number of tweets for each airline \n ",df.groupby('airline')['airline_sentime
nt'].count().sort_values(ascending=False))
airlines= ['US Airways','United','American','Southwest','Delta','Virgin America']
plt.figure(1,figsize=(12, 12))
for i in airlines:
    indices= airlines.index(i)
    plt.subplot(2,3,indices+1)
    new_df=df[df['airline']==i]
    count=new_df['airline_sentiment'].value_counts()
    Index = [1,2,3]
    plt.bar(Index,count, color=['orangered', 'lemonchiffon', 'yellowgreen'])
    plt.xticks(Index,['negative','neutral','positive'])
    plt.ylabel('Mood Count')
    plt.xlabel('Mood')
    plt.title('Count of Moods of '+i)
```

```
Total number of tweets for each airline
  airline
United            3822
US Airways        2913
American          2759
Southwest         2420
Delta             2222
Virgin America     504
Name: airline_sentiment, dtype: int64
```

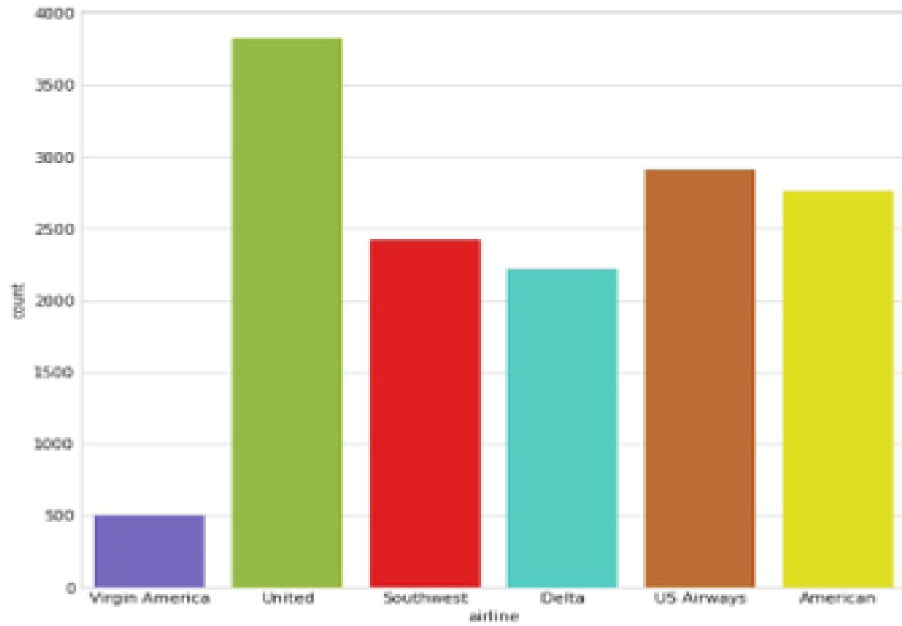




In [53]:

```
# Unique values of airline

col = ['slateblue', 'yellowgreen', 'red', 'turquoise', 'chocolate', 'yellow']
plt.figure(figsize=(9,8))
ax = sns.countplot(x="airline", data=df, palette = col)
```



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.