PROJECT DOCUMENTATION
SENTIMENT ANALYSIS MARKETING
PHASE 5

PROBLEM STATEMENT

The problem at hand involves performing sentiment analysis on

customer feedback to gain valuable 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

leveraging various Natural Language Processing (NLP) methods

to extract meaningful insights from customer feedback.

SENTIMENT ANALYSIS TECHNIQUES

 Employ different NLP techniques like Bag of Words, Word

Embeddings, or Transformer models for sentiment analysis.

Create a BoW representation of the text data, which counts the

frequency of words in each document. Utilize pre-trained word

embeddings to capture semantic meaning and relationships

between words. Leverage advanced transformer-based models

for deep contextualized sentiment analysis

AUTOMATED RESPONSE SYSTEM

Develop an automatic response system that suggests appropriate response customer feedback, especially for negative sentiment improve customer service and satisfaction.

PREDICTIVE ANALYSIS

Implement predictive analytic forecast future sentiment trends based on historical can help airlines proactively address potential issues.

FEEDBACK LOOP

Continuously improve the sentiment analysis system based on

user feedback and changing customer sentiments. Collect

feedback from users and business stakeholders regarding the

accuracy and usefulness of the sentiment analysis. Make

periodic updates to the model based on feedback and changing

customer sentiments to maintain its relevance and accuracy.

MODEL EVALUATION

Assess the performance of the sentiment analysis models and

understand how well they predict sentiment. Use appropriate

evaluation metrics (e.g., accuracy, precision, recall, F1-score) to

understand how well they predict sentiment. Use appropriate

evaluation metrics (e.g., accuracy, precision, recall, F1-score) to

measure the model's performance. Choose metrics that are

relevant to your specific business objectives. Employ crossvalidation techniques (e.g., k-fold cross-validation) to ensure the

model's generalization and robustness. Analyze the confusion

matrix to understand where the model is making errors (e.g.,

false positives, false negatives).

program

```
import pandas as pd

import numpy as np

import matplotlib.pyplot as plt

import seaborn as sns

import warnings

warnings.filterwarnings("ignore")

import os

for dirname, _, filenames in os.walk('/kaggle/input'):
 for filename in filenames:
  print(os.path.join(dirname, filename))
```

/kaggle/input/twitter-airline-sentiment/Tweets.csv

/kaggle/input/twitter-airline-sentiment/database.sqlite

In [4]:

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

In [5]:

df.head()

Out[5]:

df.info()

<class 'pandas.core.frame.DataFrame'>

RangeIndex: 14640 entries, 0 to 14639

<class 'pandas.core.frame.DataFrame'>

RangeIndex: 14640 entries, 0 to 14639

Data columns (total 10 columns):

Column Non-Null Count Dtype

```
---  ------  --------------  -----
```

0 tweet_id   14640   non-null
int64

1 airline_sentiment 14640   non-null
object

2 airline_sentiment_confidence 14640   non-null
float64

3 negativereason 9178   non-null
object

4 negativereason_confidence 10522   non-null
float64

5 airline 14640   non-null
object

6 airline_sentiment_gold 40   non-null
object

7 name 14640   non-null
object

8 negativereason_gold 32   non-null
object

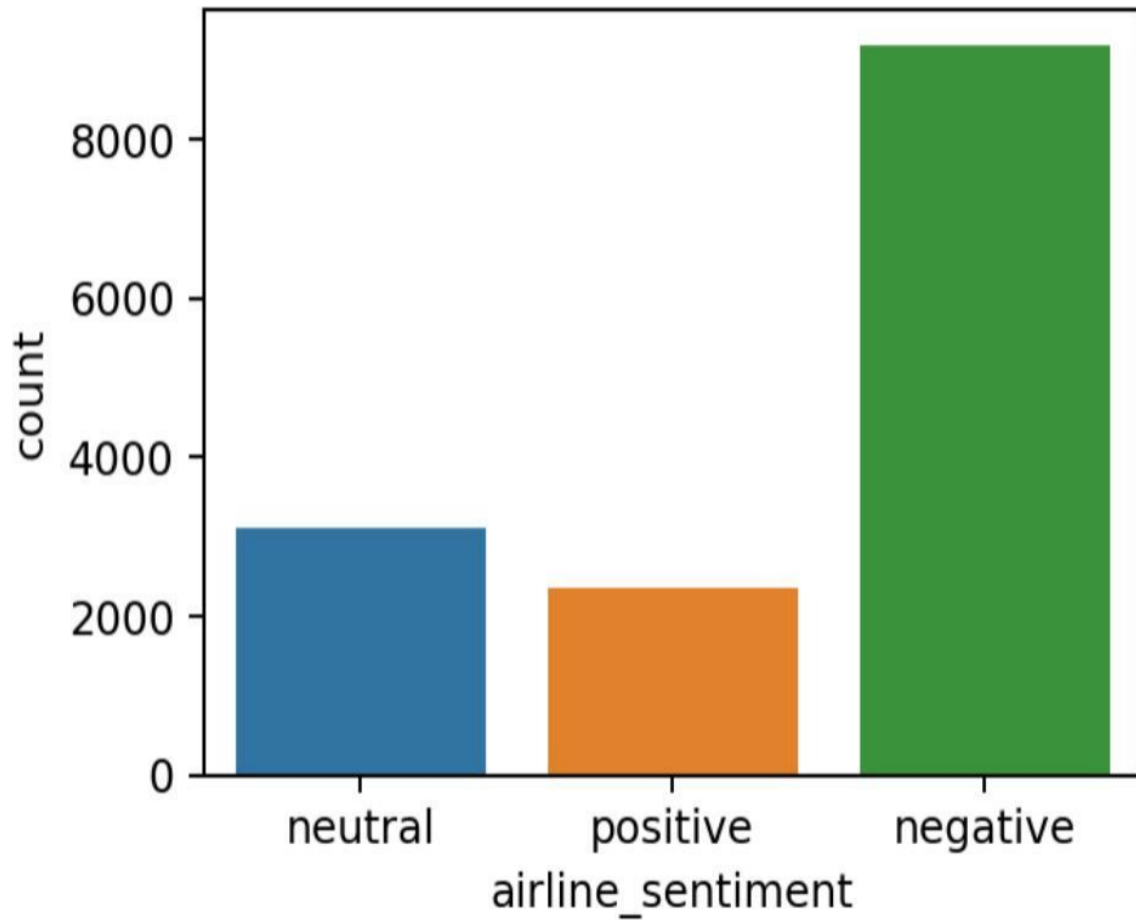9 retweet_count 14640   non-null
int64

10 text 14640   non-null

object

Exploratory Data Analysis

CONCLUSION

After performing cross validation and hyper parameter

tuning via grid

search, also evaluating the final 10 models to unseen dataset, here

are some conclusion.

• The best model is SVC (C=1, gamma=1, kernel='linear') with

76.8% accuracy and 77.5% f1 score.

• Decision Tree improves significantly after used on Bagging Classifier.

• Neutral is the hardest class label to predict

Accurately

This design thinking approach outlines a structured methodology f

or tackling the problem of

preforming sentiment analysis on customer

feedback a gain insights into competitor products.

By allowing these steps, we aim to extract valuable

information from textual data, visualize trends, and generate actionable insights that can drive informed business decisions and enhance the company's competitive edge in the market. By infusing innovation into design thinking approach, can createa powerful solution for performing sentiment

analysis on customer in the airline industry,

Ultimately leading to improved products and

services, enhanced customer satisfaction, and a

Competitive edge in the market. By following these steps, we will have a clean and loaded dataset ready for sentiment analysis. From there, we can proceed with applying sentiment analysis technique, feature extraction, visualization, and insights.