

```
%pip install pandas numpy matplotlib seaborn scikit-learn nltk  
vaderSentiment
```

Requirement already satisfied: pandas in c:\users\yasha\appdata\local\programs\python\python39\lib\site-packages (2.3.1)

Requirement already satisfied: numpy in c:\users\yasha\appdata\local\programs\python\python39\lib\site-packages (2.0.2)

Requirement already satisfied: matplotlib in c:\users\yasha\appdata\local\programs\python\python39\lib\site-packages (3.9.4)

Requirement already satisfied: seaborn in c:\users\yasha\appdata\local\programs\python\python39\lib\site-packages (0.13.2)

Requirement already satisfied: scikit-learn in c:\users\yasha\appdata\local\programs\python\python39\lib\site-packages (1.6.1)

Requirement already satisfied: nltk in c:\users\yasha\appdata\local\programs\python\python39\lib\site-packages (3.9.1)

Requirement already satisfied: vaderSentiment in c:\users\yasha\appdata\local\programs\python\python39\lib\site-packages (3.3.2)

Requirement already satisfied: tzdata>=2022.7 in c:\users\yasha\appdata\local\programs\python\python39\lib\site-packages (from pandas) (2025.2)

Requirement already satisfied: pytz>=2020.1 in c:\users\yasha\appdata\local\programs\python\python39\lib\site-packages (from pandas) (2025.2)

Requirement already satisfied: python-dateutil>=2.8.2 in c:\users\yasha\appdata\local\programs\python\python39\lib\site-packages (from pandas) (2.9.0.post0)

Requirement already satisfied: pillow>=8 in c:\users\yasha\appdata\local\programs\python\python39\lib\site-packages (from matplotlib) (11.1.0)

Requirement already satisfied: importlib-resources>=3.2.0 in c:\users\yasha\appdata\local\programs\python\python39\lib\site-packages (from matplotlib) (6.5.2)

Requirement already satisfied: pyparsing>=2.3.1 in c:\users\yasha\appdata\local\programs\python\python39\lib\site-packages (from matplotlib) (3.2.3)

Requirement already satisfied: contourpy>=1.0.1 in c:\users\yasha\appdata\local\programs\python\python39\lib\site-packages (from matplotlib) (1.3.0)

Requirement already satisfied: cycler>=0.10 in c:\users\yasha\appdata\local\programs\python\python39\lib\site-packages (from matplotlib) (0.12.1)

Requirement already satisfied: kiwisolver>=1.3.1 in c:\users\yasha\appdata\local\programs\python\python39\lib\site-packages (from matplotlib) (1.4.7)

Requirement already satisfied: packaging>=20.0 in c:\users\yasha\appdata\local\programs\python\python39\lib\site-packages (from matplotlib) (24.2)

Requirement already satisfied: fonttools>=4.22.0 in c:\users\yasha\appdata\local\programs\python\python39\lib\site-packages (from matplotlib) (4.56.0)

Requirement already satisfied: threadpoolctl>=3.1.0 in c:\users\yasha\appdata\local\programs\python\python39\lib\site-packages (from scikit-learn) (3.6.0)

Requirement already satisfied: joblib>=1.2.0 in c:\users\yasha\appdata\local\programs\python\python39\lib\site-packages (from scikit-learn) (1.4.2)

Requirement already satisfied: scipy>=1.6.0 in c:\users\yasha\appdata\local\programs\python\python39\lib\site-packages (from scikit-learn) (1.13.1)

Requirement already satisfied: tqdm in c:\users\yasha\appdata\local\programs\python\python39\lib\site-packages (from nltk) (4.67.1)

Requirement already satisfied: click in c:\users\yasha\appdata\local\programs\python\python39\lib\site-packages (from nltk) (8.1.8)

Requirement already satisfied: regex>=2021.8.3 in c:\users\yasha\appdata\local\programs\python\python39\lib\site-packages (from nltk) (2025.7.34)

Requirement already satisfied: requests in c:\users\yasha\appdata\local\programs\python\python39\lib\site-packages (from vaderSentiment) (2.32.3)

Requirement already satisfied: zipp>=3.1.0 in c:\users\yasha\appdata\local\programs\python\python39\lib\site-packages (from importlib-resources>=3.2.0->matplotlib) (3.21.0)

Requirement already satisfied: six>=1.5 in c:\users\yasha\appdata\local\programs\python\python39\lib\site-packages (from python-dateutil>=2.8.2->pandas) (1.17.0)

Requirement already satisfied: colorama in c:\users\yasha\appdata\local\programs\python\python39\lib\site-packages (from click->nltk) (0.4.6)

Requirement already satisfied: charset-normalizer<4,>=2 in c:\users\yasha\appdata\local\programs\python\python39\lib\site-packages (from requests->vaderSentiment) (3.4.1)

Requirement already satisfied: certifi>=2017.4.17 in c:\users\yasha\appdata\local\programs\python\python39\lib\site-packages (from requests->vaderSentiment) (2025.1.31)

Requirement already satisfied: idna<4,>=2.5 in c:\users\yasha\appdata\local\programs\python\python39\lib\site-packages (from requests->vaderSentiment) (3.10)

Requirement already satisfied: urllib3<3,>=1.21.1 in c:\users\yasha\appdata\local\programs\python\python39\lib\site-packages (from requests->vaderSentiment) (2.3.0)

Note: you may need to restart the kernel to use updated packages.

WARNING: You are using pip version 22.0.4; however, version 25.2 is available.

You should consider upgrading via the 'c:\Users\yasha\AppData\Local\Programs\Python\Python39\python.exe -m pip install --upgrade pip' command.

```
import pandas as pd
import numpy as np
```

```

import matplotlib.pyplot as plt
import seaborn as sns

from sklearn.model_selection import train_test_split
from sklearn.feature_extraction.text import TfidfVectorizer
from sklearn.linear_model import LogisticRegression
from sklearn.metrics import accuracy_score, confusion_matrix,
classification_report

import nltk
from nltk.corpus import stopwords
from nltk.stem import PorterStemmer
import re

nltk.download('stopwords')

[nltk_data] Downloading package stopwords to
[nltk_data] C:\Users\yasha\AppData\Roaming\nltk_data...
[nltk_data] Package stopwords is already up-to-date!

True

df =
pd.read_csv("https://raw.githubusercontent.com/DD2405/Twitter_Sentimen
t_Analysis/master/train.csv")
df = df[['tweet', 'label']]
df.columns = ['text', 'sentiment']
df.head()

      text  sentiment
0  @user when a father is dysfunctional and is s...      0
1  @user @user thanks for #lyft credit i can't us...      0
2                bihday your majesty                  0
3  #model i love u take with u all the time in ...      0
4      factsguide: society now #motivation            0

from tqdm import tqdm
tqdm.pandas()

ps = PorterStemmer()
stop_words = set(stopwords.words('english'))

def preprocess_text(text):
    text = re.sub('[^a-zA-Z]', ' ', text)
    text = text.lower()
    words = text.split()
    words = [ps.stem(word) for word in words if word not in
stop_words]
    return ' '.join(words)

corpus = df['text'].progress_apply(preprocess_text)

```

100%|██████████| 31962/31962 [00:05<00:00, 5610.23it/s]

```
tfidf = TfidfVectorizer(max_features=5000)
```

```
X = tfidf.fit_transform(corpus).toarray()
```

```
y = df['sentiment']
```

```
X_train, X_test, y_train, y_test = train_test_split(X, y,  
test_size=0.2, random_state=42)
```

```
model = LogisticRegression()
```

```
model.fit(X_train, y_train)
```

```
LogisticRegression()
```

```
y_pred = model.predict(X_test)
```

```
print("Accuracy:", accuracy_score(y_test, y_pred))
```

```
print("\nConfusion Matrix:\n", confusion_matrix(y_test, y_pred))
```

```
print("\nClassification Report:\n", classification_report(y_test,  
y_pred))
```

Accuracy: 0.9507273580478649

Confusion Matrix:

```
[[5921  16]
```

```
 [ 299 157]]
```

Classification Report:

	precision	recall	f1-score	support
0	0.95	1.00	0.97	5937
1	0.91	0.34	0.50	456
accuracy			0.95	6393
macro avg	0.93	0.67	0.74	6393
weighted avg	0.95	0.95	0.94	6393

```
def predict_sentiment(text):
```

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

```
    text = text.lower().split()
```

```
    text = [ps.stem(word) for word in text if word not in  
stopwords.words('english')]
```

```
    text = ' '.join(text)
```

```
    vector = tfidf.transform([text]).toarray()
```

```
    prediction = model.predict(vector)
```

```
    return "Positive" if prediction == 1 else "Negative"
```

```
predict_sentiment("This product is amazing and works great!")
```

```
'Negative'
```