∨ *Customer Feedback Sentiment Analysis System*

PHASE 2 — Text Preprocessing (Classical NLP)

DataSet

https://drive.google.com/file/d/1WNnl6aPtwjl8Od8iIE0WhxeR533zu9Ce/view?usp=sharing

```
# Step 1: Import libraries
import pandas as pd
import numpy as np
import re
import nltk
from nltk.corpus import stopwords
from nltk.stem import PorterStemmer
nltk.download('stopwords')
```
```
[nltk_data] Downloading package stopwords to /root/nltk_data...
[nltk_data]   Package stopwords is already up-to-date!
True
```

```
# Step 2: Load dataset
df = pd.read_csv('/content/Twitter_Data.csv')
df.head()
```

	clean_text	category
0	when modi promised "minimum government maximum...	-1.0
1	talk all the nonsense and continue all the dra...	0.0
2	what did just say vote for modi welcome bjp t...	1.0
3	asking his supporters prefix chowkidar their n...	1.0
4	answer who among these the most powerful world...	1.0

```
# Step 3: Rename columns for clarity
df = df[['clean_text', 'category']]
df.columns = ['text', 'sentiment']
```

```
print(df.isnull().sum())
```
```
text        4
sentiment   7
dtype: int64
```

```
# Step 4: Handle missing values
df.dropna(inplace=True)
```

```
print(df.isnull().sum())
```
```
text        0
sentiment   0
dtype: int64
```

```
# Step 5: Text cleaning function
stemmer = PorterStemmer()
stop_words = set(stopwords.words('english'))
def preprocess_text(text):
  text = text.lower()
  text = re.sub(r'http\S+|www\S+', '', text)
  text = re.sub(r'[^a-zA-Z]', ' ', text)
  words = text.split()
  words = [stemmer.stem(word) for word in words if word not in stop_words]
  return ' '.join(words)


# Apply preprocessing
df['clean_text'] = df['text'].apply(preprocess_text)
```

```
print(df.head(5))
```

```
                                       text  sentiment  \
0  when modi promised "minimum government maximum...      -1.0
1  talk all the nonsense and continue all the dra...       0.0
2  what did just say vote for modi  welcome bjp t...       1.0
3  asking his supporters prefix chowkidar their n...       1.0
4  answer who among these the most powerful world...       1.0

                                    clean_text
0  modi promis minimum govern maximum govern expe...
1                 talk nonsens continu drama vote modi
2  say vote modi welcom bjp told rahul main campa...
3  ask support prefix chowkidar name modi great s...
4  answer among power world leader today trump pu...
```

PHASE 3 — TF-IDF Based Model

```python
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 classification_report

# Train-test split
X_train, X_test, y_train, y_test = train_test_split(
df['clean_text'], df['sentiment'], test_size=0.2, random_state=42)

# TF-IDF Vectorization
tfidf = TfidfVectorizer(max_features=5000)
X_train_tfidf = tfidf.fit_transform(X_train)
X_test_tfidf = tfidf.transform(X_test)

# Train Logistic Regression model
lr = LogisticRegression(max_iter=1000)
lr.fit(X_train_tfidf, y_train)


# Evaluation
y_pred = lr.predict(X_test_tfidf)
print(classification_report(y_test, y_pred))
```

```
              precision    recall  f1-score   support

        -1.0       0.82      0.72      0.77      7152
         0.0       0.82      0.91      0.86     11067
         1.0       0.88      0.85      0.86     14375

    accuracy                           0.84     32594
   macro avg       0.84      0.83      0.83     32594
weighted avg       0.84      0.84      0.84     32594
```

PHASE 4 — Word2Vec Based Model

```python
!pip install gensim
```

```
Requirement already satisfied: gensim in /usr/local/lib/python3.12/dist-packages (4.4.0)
Requirement already satisfied: numpy>=1.18.5 in /usr/local/lib/python3.12/dist-packages (from gensim) (2.0.2)
Requirement already satisfied: scipy>=1.7.0 in /usr/local/lib/python3.12/dist-packages (from gensim) (1.16.3)
Requirement already satisfied: smart_open>=1.8.1 in /usr/local/lib/python3.12/dist-packages (from gensim) (7.5.0)
Requirement already satisfied: wrapt in /usr/local/lib/python3.12/dist-packages (from smart_open>=1.8.1->gensim) (2.0.1)
```

```python
from gensim.models import Word2Vec

# Tokenize sentences
tokenized_sentences = [text.split() for text in df['clean_text']]

# Train Word2Vec model
w2v_model = Word2Vec(sentences=tokenized_sentences, vector_size=100, window=5,
min_count=2)

# Sentence vector creation
def sentence_vector(sentence):
  vectors = [w2v_model.wv[word] for word in sentence.split() if word in w2v_model.wv]
  return np.mean(vectors, axis=0) if len(vectors) > 0 else np.zeros(100)

X_w2v = np.array([sentence_vector(text) for text in df['clean_text']])

# Train-test split
X_train, X_test, y_train, y_test = train_test_split(X_w2v, df['sentiment'],
test_size=0.2)
lr.fit(X_train, y_train)
```

```
y_pred = lr.predict(X_test)
print(classification_report(y_test, y_pred))
```

```
              precision    recall  f1-score   support

        -1.0       0.50      0.28      0.36      6980
         0.0       0.62      0.62      0.62     11110
         1.0       0.62      0.75      0.68     14504

    accuracy                           0.61     32594
   macro avg       0.58      0.55      0.55     32594
weighted avg       0.59      0.61      0.59     32594
```

PHASE 5 — BERT Upgrade (ADVANCED)

```
!pip install transformers
```

```
Requirement already satisfied: transformers in /usr/local/lib/python3.12/dist-packages (4.57.3)
Requirement already satisfied: filelock in /usr/local/lib/python3.12/dist-packages (from transformers) (3.20.2)
Requirement already satisfied: huggingface-hub<1.0,>=0.34.0 in /usr/local/lib/python3.12/dist-packages (from transformers) (
Requirement already satisfied: numpy>=1.17 in /usr/local/lib/python3.12/dist-packages (from transformers) (2.0.2)
Requirement already satisfied: packaging>=20.0 in /usr/local/lib/python3.12/dist-packages (from transformers) (25.0)
Requirement already satisfied: pyyaml>=5.1 in /usr/local/lib/python3.12/dist-packages (from transformers) (6.0.3)
Requirement already satisfied: regex!=2019.12.17 in /usr/local/lib/python3.12/dist-packages (from transformers) (2025.11.3)
Requirement already satisfied: requests in /usr/local/lib/python3.12/dist-packages (from transformers) (2.32.4)
Requirement already satisfied: tokenizers<=0.23.0,>=0.22.0 in /usr/local/lib/python3.12/dist-packages (from transformers) (0
Requirement already satisfied: safetensors>=0.4.3 in /usr/local/lib/python3.12/dist-packages (from transformers) (0.7.0)
Requirement already satisfied: tqdm>=4.27 in /usr/local/lib/python3.12/dist-packages (from transformers) (4.67.1)
Requirement already satisfied: fsspec>=2023.5.0 in /usr/local/lib/python3.12/dist-packages (from huggingface-hub<1.0,>=0.34.
Requirement already satisfied: typing-extensions>=3.7.4.3 in /usr/local/lib/python3.12/dist-packages (from huggingface-hub<1
Requirement already satisfied: hf-xet<2.0.0,>=1.1.3 in /usr/local/lib/python3.12/dist-packages (from huggingface-hub<1.0,>=0
Requirement already satisfied: charset_normalizer<4,>=2 in /usr/local/lib/python3.12/dist-packages (from requests->transform
Requirement already satisfied: idna<4,>=2.5 in /usr/local/lib/python3.12/dist-packages (from requests->transformers) (3.11)
Requirement already satisfied: urllib3<3,>=1.21.1 in /usr/local/lib/python3.12/dist-packages (from requests->transformers) (
Requirement already satisfied: certifi>=2017.4.17 in /usr/local/lib/python3.12/dist-packages (from requests->transformers) (
```

```
from transformers import pipeline

bert_sentiment = pipeline('sentiment-analysis')
```

```
No model was supplied, defaulted to distilbert/distilbert-base-uncased-finetuned-sst-2-english and revision 714eb0f (https:/
Using a pipeline without specifying a model name and revision in production is not recommended.
Device set to use cpu
```

```
# Prediction example
bert_sentiment("I love this product but the battery is bad")
```

```
[{'label': 'NEGATIVE', 'score': 0.9986051917076111}]
```

```
bert_sentiment("I will kill you")
```

```
[{'label': 'NEGATIVE', 'score': 0.9264492392539978}]
```

```
bert_sentiment("I love you")
```

```
[{'label': 'POSITIVE', 'score': 0.9998656511306763}]
```

```
Start coding or generate with AI.
```

PHASE 6 — Deployment Awareness (Streamlit)

```
!pip install streamlit pyngrok transformers torch
```

```
Requirement already satisfied: streamlit in /usr/local/lib/python3.12/dist-packages (1.53.0)
Collecting pyngrok
  Downloading pyngrok-7.5.0-py3-none-any.whl.metadata (8.1 kB)
Requirement already satisfied: transformers in /usr/local/lib/python3.12/dist-packages (4.57.3)
Requirement already satisfied: torch in /usr/local/lib/python3.12/dist-packages (2.9.0+cpu)
Requirement already satisfied: altair!=5.4.0,!=5.4.1,<7,>=4.0 in /usr/local/lib/python3.12/dist-packages (from streamlit)
Requirement already satisfied: blinker<2,>=1.5.0 in /usr/local/lib/python3.12/dist-packages (from streamlit) (1.9.0)
Requirement already satisfied: cachetools<7,>=5.5 in /usr/local/lib/python3.12/dist-packages (from streamlit) (6.2.4)
Requirement already satisfied: click<9,>=7.0 in /usr/local/lib/python3.12/dist-packages (from streamlit) (8.3.1)
Requirement already satisfied: numpy<3,>=1.23 in /usr/local/lib/python3.12/dist-packages (from streamlit) (2.0.2)
Requirement already satisfied: packaging>=20 in /usr/local/lib/python3.12/dist-packages (from streamlit) (25.0)
Requirement already satisfied: pandas<3,>=1.4.0 in /usr/local/lib/python3.12/dist-packages (from streamlit) (2.2.2)
Requirement already satisfied: pillow<13,>=7.1.0 in /usr/local/lib/python3.12/dist-packages (from streamlit) (11.3.0)
Requirement already satisfied: protobuf<7,>=3.20 in /usr/local/lib/python3.12/dist-packages (from streamlit) (5.29.5)
```

```
Requirement already satisfied: pyarrow>=7.0 in /usr/local/lib/python3.12/dist-packages (from streamlit) (18.1.0)
Requirement already satisfied: requests<3,>=2.27 in /usr/local/lib/python3.12/dist-packages (from streamlit) (2.32.4)
Requirement already satisfied: tenacity<10,>=8.1.0 in /usr/local/lib/python3.12/dist-packages (from streamlit) (9.1.2)
Requirement already satisfied: toml<2,>=0.10.1 in /usr/local/lib/python3.12/dist-packages (from streamlit) (0.10.2)
Requirement already satisfied: typing-extensions<5,>=4.10.0 in /usr/local/lib/python3.12/dist-packages (from streamlit) (4
Requirement already satisfied: watchdog<7,>=2.1.5 in /usr/local/lib/python3.12/dist-packages (from streamlit) (6.0.0)
Requirement already satisfied: gitpython!=3.1.19,<4,>=3.0.7 in /usr/local/lib/python3.12/dist-packages (from streamlit) (3
Requirement already satisfied: pydeck<1,>=0.8.0b4 in /usr/local/lib/python3.12/dist-packages (from streamlit) (0.9.1)
Requirement already satisfied: tornado!=6.5.0,<7,>=6.0.3 in /usr/local/lib/python3.12/dist-packages (from streamlit) (6.5.
Requirement already satisfied: PyYAML>=5.1 in /usr/local/lib/python3.12/dist-packages (from pyngrok) (6.0.3)
Requirement already satisfied: filelock in /usr/local/lib/python3.12/dist-packages (from transformers) (3.20.2)
Requirement already satisfied: huggingface-hub<1.0,>=0.34.0 in /usr/local/lib/python3.12/dist-packages (from transformers)
Requirement already satisfied: regex!=2019.12.17 in /usr/local/lib/python3.12/dist-packages (from transformers) (2025.11.3
Requirement already satisfied: tokenizers<=0.23.0,>=0.22.0 in /usr/local/lib/python3.12/dist-packages (from transformers)
Requirement already satisfied: safetensors>=0.4.3 in /usr/local/lib/python3.12/dist-packages (from transformers) (0.7.0)
Requirement already satisfied: tqdm>=4.27 in /usr/local/lib/python3.12/dist-packages (from transformers) (4.67.1)
Requirement already satisfied: setuptools in /usr/local/lib/python3.12/dist-packages (from torch) (75.2.0)
Requirement already satisfied: sympy>=1.13.3 in /usr/local/lib/python3.12/dist-packages (from torch) (1.14.0)
Requirement already satisfied: networkx>=2.5.1 in /usr/local/lib/python3.12/dist-packages (from torch) (3.6.1)
Requirement already satisfied: jinja2 in /usr/local/lib/python3.12/dist-packages (from torch) (3.1.6)
Requirement already satisfied: fsspec>=0.8.5 in /usr/local/lib/python3.12/dist-packages (from torch) (2025.3.0)
Requirement already satisfied: jsonschema>=3.0 in /usr/local/lib/python3.12/dist-packages (from altair!=5.4.0,!=5.4.1,<7,>
Requirement already satisfied: narwhals>=1.14.2 in /usr/local/lib/python3.12/dist-packages (from altair!=5.4.0,!=5.4.1,<7,
Requirement already satisfied: gitdb<5,>=4.0.1 in /usr/local/lib/python3.12/dist-packages (from gitpython!=3.1.19,<4,>=3.0
Requirement already satisfied: hf-xet<2.0.0,>=1.1.3 in /usr/local/lib/python3.12/dist-packages (from huggingface-hub<1.0,>
Requirement already satisfied: python-dateutil>=2.8.2 in /usr/local/lib/python3.12/dist-packages (from pandas<3,>=1.4.0->s
Requirement already satisfied: pytz>=2020.1 in /usr/local/lib/python3.12/dist-packages (from pandas<3,>=1.4.0->streamlit)
Requirement already satisfied: tzdata>=2022.7 in /usr/local/lib/python3.12/dist-packages (from pandas<3,>=1.4.0->streamlit
Requirement already satisfied: MarkupSafe>=2.0 in /usr/local/lib/python3.12/dist-packages (from jinja2->torch) (3.0.3)
Requirement already satisfied: charset_normalizer<4,>=2 in /usr/local/lib/python3.12/dist-packages (from requests<3,>=2.27
Requirement already satisfied: idna<4,>=2.5 in /usr/local/lib/python3.12/dist-packages (from requests<3,>=2.27->streamlit)
Requirement already satisfied: urllib3<3,>=1.21.1 in /usr/local/lib/python3.12/dist-packages (from requests<3,>=2.27->stre
Requirement already satisfied: certifi>=2017.4.17 in /usr/local/lib/python3.12/dist-packages (from requests<3,>=2.27->stre
Requirement already satisfied: mpmath<1.4,>=1.1.0 in /usr/local/lib/python3.12/dist-packages (from sympy>=1.13.3->torch) (
Requirement already satisfied: smmap<6,>=3.0.1 in /usr/local/lib/python3.12/dist-packages (from gitdb<5,>=4.0.1->gitpython
Requirement already satisfied: attrs>=22.2.0 in /usr/local/lib/python3.12/dist-packages (from jsonschema>=3.0->altair!=5.4
Requirement already satisfied: jsonschema-specifications>=2023.03.6 in /usr/local/lib/python3.12/dist-packages (from jsons
Requirement already satisfied: referencing>=0.28.4 in /usr/local/lib/python3.12/dist-packages (from jsonschema>=3.0->altai
Requirement already satisfied: rpds-py>=0.25.0 in /usr/local/lib/python3.12/dist-packages (from jsonschema>=3.0->altair!=5
Requirement already satisfied: six>=1.5 in /usr/local/lib/python3.12/dist-packages (from python-dateutil>=2.8.2->pandas<3,
Downloading pyngrok-7.5.0-py3-none-any.whl (24 kB)
Installing collected packages: pyngrok
Successfully installed pyngrok-7.5.0
```

```python
%%writefile app.py
import streamlit as st
from transformers import pipeline

st.set_page_config(page_title="Sentiment Analyzer")

st.title("Customer Feedback Sentiment Analysis")

@st.cache_resource
def load_model():
    return pipeline("sentiment-analysis")

model = load_model()

text = st.text_area("Enter customer feedback")

if st.button("Analyze"):
    if text.strip() == "":
        st.warning("Please enter text")
    else:
        result = model(text)[0]
        st.success(f"Sentiment: {result['label']}")
        st.write(f"Score: {round(result['score'], 2)}")
```

```
Writing app.py
```

```python
from pyngrok import ngrok

ngrok.set_auth_token("38VwiMXdEr8TCVT1lRDC8KpKvz7_WFLrDQSVg4pxnoPNh36w")
```

```python
!streamlit run app.py &>/content/logs.txt &

#Explanation:

#& → runs in background

#logs.txt → suppresses warnings
```

```
public_url = ngrok.connect(8501)
print("Streamlit App URL:", public_url)
```

Streamlit App URL: NgrokTunnel: "https://bibulously-nonsupporting-yoko.ngrok-free.dev" -> "http://localhost:8501"

public_url = ngrok.connect(8501)
print("Streamlit App URL:", public_url)

Streamlit App URL: NgrokTunnel: "https://bibulously-nonsupporting-yoko.ngrok-free.dev" -> "http://localhost:8501"