Load dataset

In [1]:
```python
import pandas as pd
df=pd.read_csv("C:/Users/FamiAmal/Downloads/nlp_dataset.csv")
df
```

Out[1]:

	Comment	Emotion
0	i seriously hate one subject to death but now ...	fear
1	im so full of life i feel appalled	anger
2	i sit here to write i start to dig out my feel...	fear
3	ive been really angry with r and i feel like a...	joy
4	i feel suspicious if there is no one outside l...	fear
...	...	...
5932	i begun to feel distressed for you	fear
5933	i left feeling annoyed and angry thinking that...	anger
5934	i were to ever get married i d have everything...	joy
5935	i feel reluctant in applying there because i w...	fear
5936	i just wanted to apologize to you because i fe...	anger

5937 rows × 2 columns

In [2]:
```python
df.duplicated().sum()
```

Out[2]: 0

In [3]:
```python
df.isnull()
```

Out[3]:

	Comment	Emotion
0	False	False
1	False	False
2	False	False
3	False	False
4	False	False
...	...	...
5932	False	False
5933	False	False
5934	False	False
5935	False	False
5936	False	False

5937 rows × 2 columns

In [4]:
```python
df.isnull().sum()
```

Out[4]:
```
Comment    0
Emotion    0
dtype: int64
```

Text Cleaning

```
In [5]:  import re
         def clean_text(text):

             text = re.sub(r'http\S+|www\S+|https\S+', '', text, flags=re.MULTILINE

             text = re.sub(r'[^\w\s]', '', text)

             text = text.lower()
             return text



         df['Cleaned_Text'] = df['Comment'].apply(clean_text)
         print(df['Cleaned_Text'])
```

```
0        i seriously hate one subject to death but now ...
1                       im so full of life i feel appalled
2        i sit here to write i start to dig out my feel...
3        ive been really angry with r and i feel like a...
4        i feel suspicious if there is no one outside l...
                              ...
5932                     i begun to feel distressed for you
5933     i left feeling annoyed and angry thinking that...
5934     i were to ever get married i d have everything...
5935     i feel reluctant in applying there because i w...
5936     i just wanted to apologize to you because i fe...
Name: Cleaned_Text, Length: 5937, dtype: object
```

Tokenization

```
In [6]:  import nltk
         nltk.download('punkt')
```

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

Out[6]: True

```
In [23]:  from nltk.tokenize import word_tokenize

          df["tokens"] = df["Cleaned_Text"].apply(nltk.word_tokenize)
          df["tokens"]
```

```
Out[23]:  0        [i, seriously, hate, one, subject, to, death, ...
          1               [im, so, full, of, life, i, feel, appalled]
          2        [i, sit, here, to, write, i, start, to, dig, o...
          3        [ive, been, really, angry, with, r, and, i, fe...
          4        [i, feel, suspicious, if, there, is, no, one, ...
                                    ...
          5932           [i, begun, to, feel, distressed, for, you]
          5933     [i, left, feeling, annoyed, and, angry, thinki...
          5934     [i, were, to, ever, get, married, i, d, have, ...
          5935     [i, feel, reluctant, in, applying, there, beca...
          5936     [i, just, wanted, to, apologize, to, you, beca...
          Name: tokens, Length: 5937, dtype: object
```

Stop words Removal

In [8]:
```python
from nltk.corpus import stopwords
nltk.download('stopwords')
```

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

Out[8]: True

In [9]:
```python
stop_words= set (stopwords.words('english'))
stop_words
```

Out[9]:
```
{'a',
 'about',
 'above',
 'after',
 'again',
 'against',
 'ain',
 'all',
 'am',
 'an',
 'and',
 'any',
 'are',
 'aren',
 "aren't",
 'as',
 'at',
 'be',
 'because',
 'been',
```

In [10]:
```python
df['filtered_tokens']=df['tokens'].apply(lambda x: ' '.join([word for word
df['filtered_tokens']
```

Out[10]:
```
0       seriously hate one subject death feel reluctan...
1                          im full life feel appalled
2       sit write start dig feelings think afraid acce...
3       ive really angry r feel like idiot trusting fi...
4       feel suspicious one outside like rapture happe...
                              ...
5932                        begun feel distressed
5933    left feeling annoyed angry thinking center stu...
5934    ever get married everything ready offer got to...
5935    feel reluctant applying want able find company...
5936            wanted apologize feel like heartless bitch
Name: filtered_tokens, Length: 5937, dtype: object
```

Lemmatization

In [11]:
```python
from nltk.stem import WordNetLemmatizer
lemmatizer= WordNetLemmatizer()
import spacy
nltk.download('wordnet')
nltk.download('omw-1.4')
```

```
[nltk_data] Downloading package wordnet to
[nltk_data]     C:\Users\FamiAmal\AppData\Roaming\nltk_data...
[nltk_data]   Package wordnet is already up-to-date!
[nltk_data] Downloading package omw-1.4 to
[nltk_data]     C:\Users\FamiAmal\AppData\Roaming\nltk_data...
[nltk_data]   Package omw-1.4 is already up-to-date!
```

Out[11]: True

In [14]:
```python
nlp = spacy.load("en_core_web_sm")

def lemmatize_text(text):
    doc = nlp(text)
    return ' '.join([token.lemma_ for token in doc])

df['lemmatized_tokens'] = df['filtered_tokens'].apply(lemmatize_text)
df['lemmatized_tokens']
```

```
Out[14]: 0       seriously hate one subject death feel reluctan...
         1                       I m full life feel appalled
         2       sit write start dig feeling think afraid accep...
         3       I ve really angry r feel like idiot trust firs...
         4       feel suspicious one outside like rapture happe...
                                       ...
         5932                            begin feel distressed
         5933    leave feel annoyed angry think center stupid joke
         5934    ever get marry everything ready offer get toge...
         5935    feel reluctant applying want able find company...
         5936          want apologize feel like heartless bitch
         Name: lemmatized_tokens, Length: 5937, dtype: object
```

Feature Extraction

In [15]: df

Out[15]:

	Comment	Emotion	Cleaned_Text	tokens	filtered_tokens	lemmatized_tokens
0	i seriously hate one subject to death but now ...	fear	i seriously hate one subject to death but now ...	[i, seriously, hate, one, subject, to, death, ...	seriously hate one subject death feel reluctan...	seriously hate one subject death feel reluctan...
1	im so full of life i feel appalled	anger	im so full of life i feel appalled	[im, so, full, of, life, i, feel, appalled]	im full life feel appalled	I m full life feel appalled
2	i sit here to write i start to dig out my feel...	fear	i sit here to write i start to dig out my feel...	[i, sit, here, to, write, i, start, to, dig, o...	sit write start dig feelings think afraid acce...	sit write start dig feeling think afraid accep...
3	ive been really angry with r and i feel like a...	joy	ive been really angry with r and i feel like a...	[ive, been, really, angry, with, r, and, i, fe...	ive really angry r feel like idiot trusting fi...	I ve really angry r feel like idiot trust firs...
4	i feel suspicious if there is no one outside I...	fear	i feel suspicious if there is no one outside I...	[i, feel, suspicious, if, there, is, no, one, ...	feel suspicious one outside like rapture happe...	feel suspicious one outside like rapture happe...
...	...	...	...	...	...	...
5932	i begun to feel distressed for you	fear	i begun to feel distressed for you	[i, begun, to, feel, distressed, for, you]	begun feel distressed	begin feel distressed
5933	i left feeling annoyed and angry thinking that...	anger	i left feeling annoyed and angry thinking that...	[i, left, feeling, annoyed, and, angry, thinki...	left feeling annoyed angry thinking center stu...	leave feel annoyed angry think center stupid joke
5934	i were to ever get married i d have everything...	joy	i were to ever get married i d have everything...	[i, were, to, ever, get, married, i, d, have, ...	ever get married everything ready offer got to...	ever get marry everything ready offer get toge...
5935	i feel reluctant in applying there because i w...	fear	i feel reluctant in applying there because i w...	[i, feel, reluctant, in, applying, there, beca...	feel reluctant applying want able find company...	feel reluctant applying want able find company...
5936	i just wanted to apologize to you because i fe...	anger	i just wanted to apologize to you because i fe...	[i, just, wanted, to, apologize, to, you, beca...	wanted apologize feel like heartless bitch	want apologize feel like heartless bitch

5937 rows × 6 columns

CountVectorizer

In [16]:
```python
from sklearn.feature_extraction.text import CountVectorizer
vectorizer=CountVectorizer()
bow_metrix=vectorizer.fit_transform(df)
bow_metrix
```

Out[16]:
```
<6x6 sparse matrix of type '<class 'numpy.int64'>'
        with 6 stored elements in Compressed Sparse Row format>
```

In [17]:
```python
bow_array=bow_metrix.toarray()
bow_array
```

Out[17]:
```
array([[0, 1, 0, 0, 0, 0],
       [0, 0, 1, 0, 0, 0],
       [1, 0, 0, 0, 0, 0],
       [0, 0, 0, 0, 0, 1],
       [0, 0, 0, 1, 0, 0],
       [0, 0, 0, 0, 1, 0]], dtype=int64)
```

Model Development

In [18]:
```python
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LogisticRegression
from sklearn.metrics import classification_report, confusion_matrix, accur

bow_metrix = vectorizer.fit_transform(df['lemmatized_tokens'])

# Split the data into training and testing sets
X_train, X_test, y_train, y_test = train_test_split(bow_metrix, df['Emotic
```

NAIVE BAYERS

In [19]:
```python
from sklearn.naive_bayes import GaussianNB
import numpy as np
```

In [20]:
```python
model = GaussianNB()
model.fit(X_train.toarray(), y_train)
```

Out[20]:
```
▼ GaussianNB
GaussianNB()
```

In [21]:
```python
y_pred = model.predict(X_test.toarray())


print("Classification Report:\n", classification_report(y_test, y_pred))
print("Confusion Matrix:\n", confusion_matrix(y_test, y_pred))
print("Accuracy Score:\n", accuracy_score(y_test, y_pred))
```

```
Classification Report:
               precision    recall  f1-score   support

       anger       0.68      0.55      0.61       392
        fear       0.53      0.84      0.65       416
         joy       0.77      0.42      0.54       380

    accuracy                           0.61      1188
   macro avg       0.66      0.60      0.60      1188
weighted avg       0.66      0.61      0.60      1188

Confusion Matrix:
 [[215 158  19]
 [ 38 351  27]
 [ 63 159 158]]
Accuracy Score:
 0.6094276094276094
```

SUPPORT VECTOR MACHINE - SVM

In [22]:
```python
from sklearn.svm import SVC


svm_model = SVC()
svm_model.fit(X_train, y_train)
```

Out[22]:
```
▼ SVC
SVC()
```

In [24]:
```python
y_pred = svm_model.predict(X_test)
y_pred
```

Out[24]: array(['anger', 'joy', 'fear', ..., 'joy', 'fear', 'joy'], dtype=object)

In [25]:
```python
print(confusion_matrix(y_test, y_pred))
```

```
[[358  10  24]
 [ 29 368  19]
 [  6   7 367]]
```

In [26]:
```python
print(classification_report(y_test, y_pred))
```

```
              precision    recall  f1-score   support

       anger       0.91      0.91      0.91       392
        fear       0.96      0.88      0.92       416
         joy       0.90      0.97      0.93       380

    accuracy                           0.92      1188
   macro avg       0.92      0.92      0.92      1188
weighted avg       0.92      0.92      0.92      1188
```

In [27]:
```python
accuracy=accuracy_score(y_test, y_pred)

print("Accuracy: ",accuracy)
```

```
Accuracy:  0.92003367003367
```

Model Comparison

NAIVE BAYERS f1-score

anger 0.61 fear 0.65 joy 0.54 Accuracy : 0.61

SUPPORT VECTOR MACHINE

f1-score

anger 0.91 fear 0.92 joy 0.93 Accuracy : 0.92

In [ ]: