```
import numpy as np
import pandas as pd

from google.colab import drive

drive.mount("/content/gdrive")

    Mounted at /content/gdrive

df=pd.read_csv("/content/gdrive/MyDrive/Colab Notebooks/stress.csv")
df.head()

df.describe()

df.isnull().sum()

import nltk
import re
from nltk. corpus import stopwords
import string
nltk. download( 'stopwords' )
stemmer = nltk. SnowballStemmer("english")
stopword=set (stopwords . words ( 'english' ))

def clean(text):
    text = str(text) . lower()
    text = re. sub('\[.*?\]',' ',text)
    text = re. sub('https?://\S+/www\. \S+', ' ', text)
    text = re. sub('<. *?>+', ' ', text)
    text = re. sub(' [%s]' % re. escape(string. punctuation), ' ', text)
    text = re. sub(' \n',' ', text)
    text = re. sub(' \w*\d\w*' ,' ', text)
    text = [word for word in text. split(' ') if word not in stopword]
    text =" ". join(text)
    text = [stemmer . stem(word) for word in text. split(' ') ]
    text = " ". join(text)
    return text
df [ "text"] = df["text"]. apply(clean)

import matplotlib. pyplot as plt
from wordcloud import WordCloud, STOPWORDS, ImageColorGenerator
text = " ". join(i for i in df. text)
stopwords = set (STOPWORDS)
wordcloud = WordCloud( stopwords=stopwords,background_color="white") . generate(text)
plt. figure(figsize=(10, 10) )
plt. imshow(wordcloud )
plt. axis("off")
plt. show( )

from sklearn. feature_extraction. text import CountVectorizer
from sklearn. model_selection import train_test_split

x = np.array (df["text"])
y = np.array (df["label"])

cv = CountVectorizer ()
X = cv. fit_transform(x)
print(X)
xtrain, xtest, ytrain, ytest = train_test_split(X, y,test_size=0.33)

from sklearn.naive_bayes import BernoulliNB
model=BernoulliNB()
model.fit(xtrain,ytrain)

user=input("Enter the text")
data=cv.transform([user]).toarray()
output=model.predict(data)
print(output)
```