Code

import pandas as pd

import numpy as np

data=pd.read\_csv("stress.csv")

data

import nltk

import re

nltk.download('stopwords')

stemmer = nltk.SnowballStemmer("english")

from nltk.corpus import stopwords

import string

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

data["text"] = data["text"].apply(clean)

import matplotlib.pyplot as plt

from wordcloud import WordCloud, STOPWORDS, ImageColorGenerator

text = " ".join(i for i in data.text)

stopwords = set(STOPWORDS)

wordcloud = WordCloud(stopwords=stopwords,

background\_color="white").generate(text)

plt.figure( figsize=(15,10))

plt.imshow(wordcloud, interpolation='bilinear')

plt.axis("off")

plt.show()

data["label"]=data["label"].map({0:"No Stress",1:"Stress"})

data=data[["text","label"]]

data

from sklearn.feature\_extraction.text import CountVectorizer

from sklearn.model\_selection import train\_test\_split

x=np.array(data["text"])

y=np.array(data["label"])

cv = CountVectorizer()

X = cv.fit\_transform(x)

xtrain,xtest,ytrain,ytest=train\_test\_split(X,y,test\_size=0.3,random\_state=42)

from sklearn.naive\_bayes import BernoulliNB

model = BernoulliNB()

model.fit(xtrain,ytrain)

BernoulliNB()

user = input("Text")

data=cv.transform([user]).toarray()

output=model.predict(data)

output

array(['No Stress'], dtype='<U9')