code-1

July 30, 2024

```
[26]: import pandas as pd
      import numpy as np
```

```
[27]: data = pd.read_csv("stress.csv")
      print(data.head())
```

```
            subreddit post_id sentence_range  \
0                ptsd  8601tu       (15, 20)
1          assistance  8lbrx9         (0, 5)
2                ptsd  9ch1zh       (15, 20)
3       relationships  7rorpp         [5, 10]
4    survivorsofabuse  9p2gbc         [0, 5]

                                                text     id  label  \
0  He said he had not felt that way before, sugge…  33181      1
1  Hey there r/assistance, Not sure if this is th…   2606      0
2  My mom then hit me with the newspaper and it s…  38816      1
3  until i met my new boyfriend, he is amazing, h…    239      1
4  October is Domestic Violence Awareness Month a…   1421      1

   confidence  social_timestamp  social_karma  syntax_ari  …  \
0         0.8        1521614353             5    1.806818  …
1         1.0        1527009817             4    9.429737  …
2         0.8        1535935605             2    7.769821  …
3         0.6        1516429555             0    2.667798  …
4         0.8        1539809005            24    7.554238  …

   lex_dal_min_pleasantness  lex_dal_min_activation  lex_dal_min_imagery  \
0                     1.000                  1.1250                  1.0
1                     1.125                  1.0000                  1.0
2                     1.000                  1.1429                  1.0
3                     1.000                  1.1250                  1.0
4                     1.000                  1.1250                  1.0

   lex_dal_avg_activation  lex_dal_avg_imagery  lex_dal_avg_pleasantness  \
0                 1.77000              1.52211                   1.89556
1                 1.69586              1.62045                   1.88919
2                 1.83088              1.58108                   1.85828
```

```
3              1.75356              1.52114                1.98848
4              1.77644              1.64872                1.81456

   social_upvote_ratio  social_num_comments  syntax_fk_grade  sentiment
0                 0.86                    1         3.253573  -0.002742
1                 0.65                    2         8.828316   0.292857
2                 0.67                    0         7.841667   0.011894
3                 0.50                    5         4.104027   0.141671
4                 1.00                    1         7.910952  -0.204167

[5 rows x 116 columns]
```

[28]: `data.shape`

[28]: (2838, 116)

[31]: `data.head(5)`

[31]:
```
          subreddit post_id sentence_range  \
0              ptsd  8601tu        (15, 20)
1        assistance  8lbrx9          (0, 5)
2              ptsd  9ch1zh        (15, 20)
3     relationships  7rorpp         [5, 10]
4  survivorsofabuse  9p2gbc          [0, 5]

                                                text     id  label  \
0  He said he had not felt that way before, sugge…  33181      1
1  Hey there r/assistance, Not sure if this is th…   2606      0
2  My mom then hit me with the newspaper and it s…  38816      1
3  until i met my new boyfriend, he is amazing, h…    239      1
4  October is Domestic Violence Awareness Month a…   1421      1

   confidence  social_timestamp  social_karma  syntax_ari  … \
0         0.8        1521614353             5    1.806818  …
1         1.0        1527009817             4    9.429737  …
2         0.8        1535935605             2    7.769821  …
3         0.6        1516429555             0    2.667798  …
4         0.8        1539809005            24    7.554238  …

   lex_dal_min_pleasantness  lex_dal_min_activation  lex_dal_min_imagery  \
0                     1.000                  1.1250                  1.0
1                     1.125                  1.0000                  1.0
2                     1.000                  1.1429                  1.0
3                     1.000                  1.1250                  1.0
4                     1.000                  1.1250                  1.0

   lex_dal_avg_activation  lex_dal_avg_imagery  lex_dal_avg_pleasantness  \
```

```
0                1.77000            1.52211                1.89556
1                1.69586            1.62045                1.88919
2                1.83088            1.58108                1.85828
3                1.75356            1.52114                1.98848
4                1.77644            1.64872                1.81456

   social_upvote_ratio  social_num_comments  syntax_fk_grade  sentiment
0                 0.86                    1         3.253573  -0.002742
1                 0.65                    2         8.828316   0.292857
2                 0.67                    0         7.841667   0.011894
3                 0.50                    5         4.104027   0.141671
4                 1.00                    1         7.910952  -0.204167

[5 rows x 116 columns]
```

[32]: `data.columns`

[32]:
```
Index(['subreddit', 'post_id', 'sentence_range', 'text', 'id', 'label',
       'confidence', 'social_timestamp', 'social_karma', 'syntax_ari',
       …
       'lex_dal_min_pleasantness', 'lex_dal_min_activation',
       'lex_dal_min_imagery', 'lex_dal_avg_activation', 'lex_dal_avg_imagery',
       'lex_dal_avg_pleasantness', 'social_upvote_ratio',
       'social_num_comments', 'syntax_fk_grade', 'sentiment'],
      dtype='object', length=116)
```

[34]: `data.info()`

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 2838 entries, 0 to 2837
Columns: 116 entries, subreddit to sentiment
dtypes: float64(106), int64(6), object(4)
memory usage: 2.5+ MB
```

[38]: `pip install textblob`

```
Collecting textblob
  Downloading textblob-0.18.0.post0-py3-none-any.whl (626 kB)
     ---------------------------------- 626.3/626.3 kB 73.2 kB/s eta 0:00:00
Collecting nltk>=3.8
  Using cached nltk-3.8.1-py3-none-any.whl (1.5 MB)
Requirement already satisfied: click in c:\users\m.pujitha\anaconda3\lib\site-
packages (from nltk>=3.8->textblob) (8.0.4)
Requirement already satisfied: tqdm in c:\users\m.pujitha\anaconda3\lib\site-
packages (from nltk>=3.8->textblob) (4.64.1)
Requirement already satisfied: regex>=2021.8.3 in
c:\users\m.pujitha\anaconda3\lib\site-packages (from nltk>=3.8->textblob)
(2022.7.9)
```

```
Requirement already satisfied: joblib in c:\users\m.pujitha\anaconda3\lib\site-
packages (from nltk>=3.8->textblob) (1.1.0)
Requirement already satisfied: colorama in
c:\users\m.pujitha\anaconda3\lib\site-packages (from click->nltk>=3.8->textblob)
(0.4.5)
Installing collected packages: nltk, textblob
  Attempting uninstall: nltk
    Found existing installation: nltk 3.7
    Uninstalling nltk-3.7:
      Successfully uninstalled nltk-3.7
Successfully installed nltk-3.8.1 textblob-0.18.0.post0
Note: you may need to restart the kernel to use updated packages.
```

[39]:
```python
from textblob import TextBlob
```

[40]:
```python
TextBlob("the best").polarity
```

[40]: 1.0

[41]:
```python
TextBlob("the best").sentiment
```

[41]: Sentiment(polarity=1.0, subjectivity=0.3)

[42]:
```python
def detect_sentiment(text):
    return TextBlob(text).sentiment.polarity
```

[45]:
```python
data=data[["text"]]
```

[47]:
```python
data.head()
```

[47]:
```
                                                text
0  He said he had not felt that way before, sugge…
1  Hey there r/assistance, Not sure if this is th…
2  My mom then hit me with the newspaper and it s…
3  until i met my new boyfriend, he is amazing, h…
4  October is Domestic Violence Awareness Month a…
```

[94]:
```python
data["sentiment"]=data["text"].apply(detect_sentiment)
```

[49]:
```python
data.head()
```

[49]:
```
                                                text   sentiment
0  He said he had not felt that way before, sugge…   -0.002742
1  Hey there r/assistance, Not sure if this is th…    0.292857
2  My mom then hit me with the newspaper and it s…    0.011894
3  until i met my new boyfriend, he is amazing, h…    0.141671
4  October is Domestic Violence Awareness Month a…   -0.204167
```

```
[51]: data.sentiment.value_counts()
```

```
[51]: 0.000000    63
      0.250000    21
      0.100000    19
      0.175000    14
      0.166667    12
                  ..
      0.038095     1
      0.032727     1
      0.166171     1
      0.095833     1
      0.016667     1
      Name: sentiment, Length: 2137, dtype: int64
```

```
[52]: import nltk
      import re
      stemmer = nltk.SnowballStemmer("english")
      from nltk.corpus import stopwords
      import string
```

```
[53]: stopwords = set(stopwords.words("english"))
```

```
[93]: 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 stopwords]
          text=" ".join(text)
          text = [stemmer.stem(word) for word in text.split(' ')]
          text=" ".join(text)
          return text
      data["text"] = data["text"].apply(clean)
```

```
[57]: data["text"]
```

```
[57]: 0        said felt way sugget go rest trigger ahead you…
      1        hey rassist sure right place post goe  im curr…
      2        mom hit newspap shock would know dont like pla…
      3        met new boyfriend amaz kind sweet good student…
      4        octob domest violenc awar month domest violenc…
                                    …
      2833     week ago precious ignor  jan  happi year prec…
```
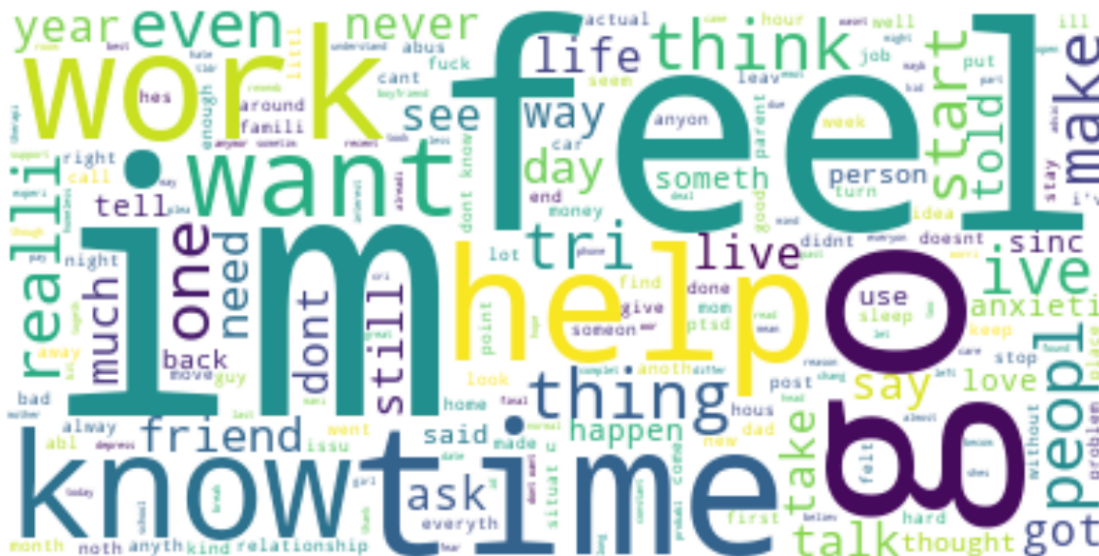
```
2834    dont abil cope anymor im tri lot thing trigger…
2835    case first time your read post look peopl will…
2836    find normal good relationship main problem see…
2837    talk mom morn said sister trauma wors mine did…
Name: text, Length: 2838, dtype: object
```

[67]:
```python
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()
```


[69]:
```python
data.head()
```

[69]:

	text	sentiment
0	said felt way sugget go rest trigger ahead you…	-0.002742
1	hey rassist sure right place post goe im curr…	0.292857
2	mom hit newspap shock would know dont like pla…	0.011894
3	met new boyfriend amaz kind sweet good student…	0.141671
4	octob domest violenc awar month domest violenc…	-0.204167

[71]:
```python
import warnings
warnings.filterwarnings("ignore")
```

```python
[72]: data["sentiment"]=data["text"].apply(detect_sentiment)
```

```python
[73]: data.head()
```

```
[73]:                                                   text   sentiment
      0  said felt way sugget go rest trigger ahead you…   0.027273
      1  hey rassist sure right place post goe  im curr…   0.340816
      2  mom hit newspap shock would know dont like pla…   0.124091
      3  met new boyfriend amaz kind sweet good student…   0.229172
      4  octob domest violenc awar month domest violenc…   0.083333
```

```python
[79]: from sklearn.feature_extraction.text import CountVectorizer, TfidfVectorizer

      from sklearn.model_selection import train_test_split
      from sklearn.naive_bayes import MultinomialNB
      from sklearn.metrics import accuracy_score
```

```python
[80]: vect=CountVectorizer(stop_words="english")
```

```python
[81]: x=vect.fit_transform(x)
```

```python
[82]: x_train,x_test,y_train,y_test=train_test_split(x,y,random_state=42)
```

```python
[83]: mb=MultinomialNB()
```

```python
[84]: tahmin=mb.fit(x_train,y_train).predict(x_test)
```

```python
[85]: accuracy_score(tahmin,y_test)
```

```
[85]: 0.7338028169014085
```

```python
[86]: from sklearn.tree import DecisionTreeClassifier
```

```python
[87]: d=DecisionTreeClassifier()
```

```python
[88]: d.fit(x_train,y_train)
```

```
[88]: DecisionTreeClassifier()
```

```python
[89]: tahmin1=d.predict(x_test)
```

```python
[90]: accuracy_score(y_test,tahmin1)
```

```
[90]: 0.6084507042253521
```

```python
[100]: user="sometimes i feel like I need some help"
```

```
[101]:  data=vect.transform([user]).toarray()
        output=d.predict(data)
        print(output)
```

['Stress']

[ ]: