

```
import nltk
nltk.download('vader_lexicon')
from nltk.sentiment.vader import SentimentIntensityAnalyzer
sia = SentimentIntensityAnalyzer()
import pandas as pd
```

```
[nltk_data] Downloading package vader_lexicon to /root/nltk_data...
[nltk_data] Package vader_lexicon is already up-to-date!
```

```
data= pd.read_csv('amazon_reviews.tsv', sep="\t") ## Tells that it is a tab seperated file
```

```
data.head()
```

```
[nltk_data] Downloading package vader_lexicon to /root/nltk_data...
[nltk_data] Package vader_lexicon is already up-to-date!
```

	label	review
0	pos	Stuning even for the non-gamer: This sound tra...
1	pos	The best soundtrack ever to anything.: I'm rea...
2	pos	Amazing!: This soundtrack is my favorite music...
3	pos	Excellent Soundtrack: I truly like this soundt...
4	pos	Remember, Pull Your Jaw Off The Floor After He...

```
len(data)
```

```
[nltk_data] Downloading package vader_lexicon to /root/nltk_data...
[nltk_data] Package vader_lexicon is already up-to-date!
```

```
2984
```

```
data['label'].unique()
```

```
[nltk_data] Downloading package vader_lexicon to /root/nltk_data...
[nltk_data] Package vader_lexicon is already up-to-date!
```

```
array(['pos', 'neg'], dtype=object)
```

```
data['label'].value_counts()
```

```
[nltk_data] Downloading package vader_lexicon to /root/nltk_data...
[nltk_data] Package vader_lexicon is already up-to-date!
```

```
neg    1574
pos    1410
Name: label, dtype: int64
```

```
## Calculating the polarity score of each review
polarity_score=[]
for i in data['review']:
    polarity_score.append(sia.polarity_scores(i))
data['Polarity Score']= polarity_score
data
```

```
[nltk_data] Downloading package vader_lexicon to /root/nltk_data...
[nltk_data] Package vader_lexicon is already up-to-date!
```

	label	review	Polarity Score
0	pos	Stuning even for the non-gamer: This sound tra...	{'neg': 0.088, 'neu': 0.669, 'pos': 0.243, 'co...
1	pos	The best soundtrack ever to anything.: I'm rea...	{'neg': 0.018, 'neu': 0.837, 'pos': 0.145, 'co...
2	pos	Amazing!: This soundtrack is my favorite music...	{'neg': 0.04, 'neu': 0.692, 'pos': 0.268, 'com...
3	pos	Excellent Soundtrack: I truly like this soundt...	{'neg': 0.09, 'neu': 0.615, 'pos': 0.295, 'com...
4	pos	Remember, Pull Your Jaw Off The Floor After He...	{'neg': 0.0, 'neu': 0.746, 'pos': 0.254, 'comp...
...
2979	neg	Not a good movie. Not for children.: Not recom...	{'neg': 0.167, 'neu': 0.833, 'pos': 0.0, 'comp...
2980	pos	movies on time: The movies get to me the same ...	{'neg': 0.0, 'neu': 1.0, 'pos': 0.0, 'compound...
2981	neg	No Special Features: First, the movie has incr...	{'neg': 0.224, 'neu': 0.616, 'pos': 0.16, 'com...
2982	neg	Horrible.: The weakest link of the Transformer...	{'neg': 0.182, 'neu': 0.818, 'pos': 0.0, 'comp...
2983	neg	Horrible film: Horrible film that should be av...	{'neg': 0.228, 'neu': 0.73, 'pos': 0.042, 'com...

```
2984 rows x 4 columns
```

```
print(data['Polarity Score'])
```

```
[nltk_data] Downloading package vader_lexicon to /root/nltk_data...
[nltk_data] Package vader_lexicon is already up-to-date!
```

```
0      {'neg': 0.088, 'neu': 0.669, 'pos': 0.243, 'co...
1      {'neg': 0.018, 'neu': 0.837, 'pos': 0.145, 'co...
2      {'neg': 0.04, 'neu': 0.692, 'pos': 0.268, 'com...
3      {'neg': 0.09, 'neu': 0.615, 'pos': 0.295, 'com...
4      {'neg': 0.0, 'neu': 0.746, 'pos': 0.254, 'comp...
...
2979   {'neg': 0.167, 'neu': 0.833, 'pos': 0.0, 'comp...
2980   {'neg': 0.0, 'neu': 1.0, 'pos': 0.0, 'compound...
2981   {'neg': 0.224, 'neu': 0.616, 'pos': 0.16, 'com...
```



McAfee WebAdvisor



Your download's being scanned.
We'll let you know if there's an issue.

```

2982     {'neg': 0.182, 'neu': 0.818, 'pos': 0.0, 'comp...
2983     {'neg': 0.228, 'neu': 0.73, 'pos': 0.042, 'com...
Name: Polarity Score, Length: 2984, dtype: object

```

```

## Predicting the sentiment (Basically its a classification problem as we need to decide on neg/pos/neu)
Sentiment=[]
for i in data['Polarity Score']:
    c= i['compound']
    if c>0:
        Sentiment.append("pos")
    else:
        Sentiment.append("neg")

```

```

data['Sentiment']= Sentiment
data.head(20)

```

	label	review	Polarity Score	Sentiment
0	pos	Stuning even for the non-gamer: This sound tra...	{'neg': 0.088, 'neu': 0.669, 'pos': 0.243, 'co...	pos
1	pos	The best soundtrack ever to anything.: I'm rea...	{'neg': 0.018, 'neu': 0.837, 'pos': 0.145, 'co...	pos
2	pos	Amazing!: This soundtrack is my favorite music...	{'neg': 0.04, 'neu': 0.692, 'pos': 0.268, 'com...	pos
3	pos	Excellent Soundtrack: I truly like this soundt...	{'neg': 0.09, 'neu': 0.615, 'pos': 0.295, 'com...	pos
4	pos	Remember, Pull Your Jaw Off The Floor After He...	{'neg': 0.0, 'neu': 0.746, 'pos': 0.254, 'comp...	pos
5	pos	an absolute masterpiece: I am quite sure any o...	{'neg': 0.014, 'neu': 0.737, 'pos': 0.249, 'co...	pos
6	neg	Buyer beware: This is a self-published book, a...	{'neg': 0.124, 'neu': 0.806, 'pos': 0.069, 'co...	neg
7	pos	Glorious story: I loved Whisper of the wicked ...	{'neg': 0.064, 'neu': 0.588, 'pos': 0.349, 'co...	pos
8	pos	A FIVE STAR BOOK: I just finished reading Whis...	{'neg': 0.113, 'neu': 0.712, 'pos': 0.174, 'co...	pos
9	pos	Whispers of the Wicked Saints: This was a easy...	{'neg': 0.033, 'neu': 0.777, 'pos': 0.19, 'com...	pos
10	neg	The Worst!: A complete waste of time. Typograp...	{'neg': 0.36, 'neu': 0.586, 'pos': 0.054, 'com...	neg
11	pos	Great book: This was a great book,I just could...	{'neg': 0.0, 'neu': 0.836, 'pos': 0.164, 'comp...	pos
12	pos	Great Read: I thought this book was brilliant,...	{'neg': 0.044, 'neu': 0.554, 'pos': 0.402, 'co...	pos
13	neg	Oh please: I guess you have to be a romance no...	{'neg': 0.077, 'neu': 0.723, 'pos': 0.199, 'co...	pos
14	neg	Awful beyond belief!: I feel I have to write t...	{'neg': 0.171, 'neu': 0.755, 'pos': 0.074, 'co...	neg
15	neg	Don't try to fool us with fake reviews.: It's ...	{'neg': 0.105, 'neu': 0.832, 'pos': 0.063, 'co...	neg
16	pos	A romantic zen baseball comedy: When you hear ...	{'neg': 0.021, 'neu': 0.818, 'pos': 0.161, 'co...	pos
17	pos	Fashionable Compression Stockings!: After I ha...	{'neg': 0.11, 'neu': 0.754, 'pos': 0.136, 'com...	pos
18	pos	Jobst UltraSheer Thigh High: Excellent product...	{'neg': 0.064, 'neu': 0.771, 'pos': 0.165, 'co...	pos
19	neg	sizes recommended in the size chart are not rea...	{'neg': 0.0, 'neu': 0.935, 'pos': 0.065, 'comp...	pos

```

## Calculating the confusion matrix
from sklearn.metrics import confusion_matrix
cm= confusion_matrix(data['label'],data['Sentiment']) ##(actual,predicted)
cm

```

```

array([[ 852,  722],
       [ 145, 1265]])

```

```

## Calculating the accuracy
from sklearn.metrics import accuracy_score
a= accuracy_score(data['label'],data['Sentiment'])*100
print("The accuracy is: "+ str(a)+"%")

```

```

The accuracy is: 70.9450402144772%

```

```

## In order to improve the accuracy we need to create our own library with adding words daily!!

```

```

## You have twitter data do the same on it
## Image processing will be done on offline python

```



McAfee WebAdvisor



Your download's being scanned.
We'll let you know if there's an issue.



McAfee WebAdvisor



Your download's being scanned.
We'll let you know if there's an issue.