

Github :- [https://github.com/gadren/ADS\\_Spring2018](https://github.com/gadren/ADS_Spring2018)

Python libraries **requests** and **lxml** to scrape reviews from an amazon product page. Rather than scraping the reviews from many products I decided to look at one product with a lot of review which happens to be the Echo Dot (2nd Generation). It counts more than 60.000 reviews.

Using this URL we will get the most recent reviews first and then we just modify the **page Number** parameter in the url to go from one page to the next. Before requesting the data from Amazon get your browser's user agent string to avoid being blocked by Amazon

```
In [*]: from lxml import html
import requests
import datetime
import pandas as pd
import numpy as np
import string

In [*]: review_df = []
xpath_rating = '//*[@data-hook="review-star-rating"]//text()'
xpath_title = '//*[@data-hook="review-title"]//text()'
xpath_author = '//*[@data-hook="review-author"]//text()'
xpath_date = '//*[@data-hook="review-date"]//text()'
xpath_body = '//*[@data-hook="review-body"]//text()'
xpath_helpful = '//*[@data-hook="helpful-vote-statement"]//text()'
for i in range(1,501):
    amazon_url = 'https://www.amazon.com/product-reviews/B01DFKC2S0?pageNumber='+str(i)
    user_agent = 'Mozilla/5.0 (Windows NT 10.0; Win64; x64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/65.0.3325.181 Safari/53
    headers = {'User-Agent': user_agent}
    page = requests.get(amazon_url, headers = headers)
    parser = html.fromstring(page.content)
    xpath_reviews = '//*[@data-hook="review"]'
    reviews = parser.xpath(xpath_reviews)
    for review in reviews:
        rating = review.xpath(xpath_rating)
        title = review.xpath(xpath_title)
        author = review.xpath(xpath_author)
        date = review.xpath(xpath_date)
        body = review.xpath(xpath_body)
        helpful = review.xpath(xpath_helpful)

        review_dict = {'rating': rating,
                        'title': title,
                        'author': author,
                        'date': date,
                        'body': body,
                        'helpful': helpful}
        review_df.append(review_dict['body'])
```

Passing the XPATH to the parser will return the list with the reviews on that page

```
In [3]: len(review_df)
```

```
Out[3]: 5000
```

As we can see from above we got 5000 Amazon Reviews

```
In [4]: print(review_df)
```

```
[["My brother Robert who has been bed ridden and paralyzed with Multiple Sclerosis from his neck down for more than 30 years now has a new friend named Alexa! He was in tears with happiness when Alexa played 70's music, played Jeopardy, answered all his questions and wakes him up every morning. Thank you Amazon for giving my brother a new bedside companion.", 'Happy Holidays', 'Roy'], [{"Having worked in the electronics retail industry for years now, I've seen scores of Smart devices come and go. Until now, nobody quite got it right. In the Echo Dot, Amazon has created a near perfect blend of hardware and software. I've seen plenty of the former, but truly seamless multi platform software has eluded everyone but Amazon. We're talking major players like Samsung and Google who have been at it for much longer than Amazon. The main problem is that excellent products like the Samsung Smart Things hub, which do a fantastic job of unifying a slew of different connected devices from different companies (Nest, Honeywell, Phillips, and so on), still lacked the web connectivity and entertainment support I wanted, so I'd still end up needing my tablet or phone. Thanks to fantastic third party support, the Dot has no problem controlling all of my smart stuff while allowing me to listen to music, order food, check the weather, listen to the radio, set alarms and timers, all of which is easily accomplished through simple voice commands. To me, this is the exceedingly rare product that I didn't know I needed, and now can't live without. Similar to the smartphone and tablet I use every day that didn't exist just a few short years ago.", "If you're even a little bit curious you owe it to yourself to give the dot a try. Add a good speaker and enjoy just how simple a connected life can be!"], "Update: After a bit more time with the dot, or maybe I should say dots since I went out and bought another one for my living room, I've come up with a few tips.", '1. Use the best speakers you can with it. I found that while Bluetooth was convenient I got much better sound out of my JBL duet computer speakers.', '2. Take the time to voice train Alexa at least once. It's kinda tedious but really improves the accuracy. I've now gone through three trainings with each dot, the phrasing gets more intricate with each, and it really is amazing how much of an improvement it makes. Kinda hard to quantify, but I'd guess Alexa is at least twice as likely to understand long, complex phrases']]
```

We will now try and clean the text in the review First we will try and eliminate the numbers from reviews

We will now try and clean the text in the review First we will try and eliminate the numbers from reviews

```
In [5]: no_int = []
        for item in review_df:
            for each in item:
                x = ''.join([i for i in each if not i.isdigit()])
                no_int.append(x)
```

```
In [6]: print(no_int)
```

```
[["My brother Robert who has been bed ridden and paralyzed with Multiple Sclerosis from his neck down for more than years now has a new friend named Alexa! He was in tears with happiness when Alexa played 's music, played Jeopardy, answered all his questions and wakes him up every morning. Thank you Amazon for giving my brother a new bedside companion.", 'Happy Holidays', 'Roy'], [{"Having worked in the electronics retail industry for years now, I've seen scores of Smart devices come and go. Until now, nobody quite got it right. In the Echo Dot, Amazon has created a near perfect blend of hardware and software. I've seen plenty of the former, but truly seamless multi platform software has eluded everyone but Amazon. We're talking major players like Samsung and Google who have been at it for much longer than Amazon. The main problem is that excellent products like the Samsung Smart Things hub, which do a fantastic job of unifying a slew of different connected devices from different companies (Nest, Honeywell, Phillips, and so on), still lacked the web connectivity and entertainment support I wanted, so I'd still end up needing my tablet or phone. Thanks to fantastic third party support, the Dot has no problem controlling all of my smart stuff while allowing me to listen to music, order food, check the weather, listen to the radio, set alarms and timers, all of which is easily accomplished through simple voice commands. To me, this is the exceedingly rare product that I didn't know I needed, and now can't live without. Similar to the smartphone and tablet I use every day that didn't exist just a few short years ago.", "If you're even a little bit curious you owe it to yourself to give the dot a try. Add a good speaker and enjoy just how simple a connected life can be!"], "Update: After a bit more time with the dot, or maybe I should say dots since I went out and bought another one for my living room, I've come up with a few tips.", '1. Use the best speakers you can with it. I found that while Bluetooth was convenient I got much better sound out of my JBL duet computer speakers.', '2. Take the time to voice train Alexa at least once. It's kinda tedious but really improves the accuracy. I've now gone through three trainings with each dot, the phrasing gets more intricate with each, and it really is amazing how much of an improvement it makes. Kinda hard to quantify, but I'd guess Alexa is at least twice as likely to understand long, complex phrases']]
```

Now we will remove all the punctuation marks from the reviews

Now we will remove all the punctuation marks from the reviews

```
In [7]: import re
import string
```

```
In [8]: no_int = ''.join(c for c in s if c not in string.punctuation) for s in no_int]
```

```
In [9]: print(no_int)
```

```
['My brother Robert who has been bed ridden and paralyzed with Multiple Sclerosis from his neck down for more than years now has a new friend named Alexa He was in tears with happiness when Alexa played s music played Jeopardy answered all his questions and wakes him up every morning Thank you Amazon for giving my brother a new bedside companion', 'Happy Holidays', 'Ro y', '\xa0Having worked in the electronics retail industry for years now Ive seen scores of Smart devices come and go Until now nobody quite got it right In the Echo Dot Amazon has created a near perfect blend of hardware and software Ive seen plenty of the former but truly seamless multi platform software has eluded everyone but Amazon Were talking major players like Samsung and Google who have been at it for much longer than Amazon The main problem is that excellent products like the Samsung Smart Things hub which do a fantastic job of unifying a slew of different connected devices from different companies Nest Honeywell Phillips and so on still lacked the web connectivity and entertainment support I wanted so Id still end up needing my tablet or phone Thanks to fantastic third party support the Dot has no problem controlling all of my smart stuff while allowing me to listen to music order food check the weather listen to the radio set alarms and timers all of which is easily accomplished through simple voice commands To me this is the exceedingly rare product that I didnt know I needed and now cant live without Similar to the smartphone and tablet I use every day that didnt exist just a few short years ago', 'If youre even a little bit curious you owe it to yourself to give the dot a try Add a good speaker and enjoy just how simple a connected life can be', 'Update After a bit more time with the dot or maybe I should say dots since I went out and bought another one for my living room Ive come up with a few tips', ' Use the best speakers you can with it I found that while Bluetooth was convenient I got much better sound out of my JBL duet computer speakers', ' Take the time to voice train Alexa at least once Its kinda tedious but really improves the accuracy Ive now gone through three trainings with each dot the phrasing gets more intricate with each and it really is amazing how much of an improvement it makes Kinda hard to quantify but Id guess Alexa is at least twice as likely to understand long complex phrases and has also gained noticeable accuracy when ambient noise I
```

In the next section I will try and remove all the emojis from the reviews

In the next section I will try and remove all the emojis from the reviews

```
In [10]: import re
emoji=[]
for item in no_int:
    emoji_pattern = re.compile("[["
        u"\U0001F600-\U0001F64F" # emoticons
        u"\U0001F300-\U0001F5FF" # symbols & pictographs
        u"\U0001F680-\U0001F6FF" # transport & map symbols
        u"\U0001F1E0-\U0001F1FF" # flags (iOS)
        "]+", flags=re.UNICODE)
    x = emoji_pattern.sub('', item)
    emoji.append(x)
```

```
In [11]: print(emoji)
```

```
['My brother Robert who has been bed ridden and paralyzed with Multiple Sclerosis from his neck down for more than years now has a new friend named Alexa He was in tears with happiness when Alexa played s music played Jeopardy answered all his questions and wakes him up every morning Thank you Amazon for giving my brother a new bedside companion', 'Happy Holidays', 'Ro y', '\xa0Having worked in the electronics retail industry for years now Ive seen scores of Smart devices come and go Until now nobody quite got it right In the Echo Dot Amazon has created a near perfect blend of hardware and software Ive seen plenty of the former but truly seamless multi platform software has eluded everyone but Amazon Were talking major players like Samsung and Google who have been at it for much longer than Amazon The main problem is that excellent products like the Samsung Smart Things hub which do a fantastic job of unifying a slew of different connected devices from different companies Nest Honeywell Phillips and so on still lacked the web connectivity and entertainment support I wanted so Id still end up needing my tablet or phone Thanks to fantastic third party support the Dot has no problem controlling all of my smart stuff while allowing me to listen to music order food check the weather listen to the radio set alarms and timers all of which is easily accomplished through simple voice commands To me this is the exceedingly rare product that I didnt know I needed and now cant live without Similar to the smartphone and tablet I use every day that didnt exist just a few short years ago', 'If youre even a little bit curious you owe it to yourself to give the dot a try Add a good speaker and enjoy just how simple a connected life can be', 'Update After a bit more time with the dot or maybe I should say dots since I went out and bought another one for my living room Ive come up with a few tips', ' Use the best speakers you can with it I found that while Bluetooth was convenient I got much better sound out of my JBL duet computer speakers', ' Take the time to voice train Alexa at least once Its kinda tedious but really improves the accuracy Ive now gone through three trainings with each dot the phrasing gets more intricate with each and it really is amazing how much of an improvement it makes Kinda hard to quantify but Id guess Alexa is at least twice as likely to understand long complex phrases and has also gained noticeable accuracy when ambient noise I
```

As we will compare all the words in reviews with the words in lexicon we will make sure all the words in review are in lower case

As we will compare all the words in reviews with the words in lexicon we will make sure all the words in review are in lower case

```
In [12]: reviews = [x.lower() for x in emoji]
         print(reviews)
```

```
['my brother robert who has been bed ridden and paralyzed with multiple sclerosis from his neck down for more than years no
w has a new friend named alexa he was in tears with happiness when alexa played s music played jeopardy answered all his que
stions and wakes him up every morning thank you amazon for giving my brother a new bedside companion', 'happy holidays', 'ro
y', '\xa0having worked in the electronics retail industry for years now ive seen scores of smart devices come and go until n
ow nobody quite got it right in the echo dot amazon has created a near perfect blend of hardware and software ive seen plen
ty of the former but truly seamless multi platform software has eluded everyone but amazon were talking major players like s
amsung and google who have been at it for much longer than amazon the main problem is that excellent products like the samsu
ng smart things hub which do a fantastic job of unifying a slew of different connected devices from different companies nest
honeywell phillips and so on still lacked the web connectivity and entertainment support i wanted so id still end up needing
my tablet or phone thanks to fantastic third party support the dot has no problem controlling all of my smart stuff while al
lowing me to listen to music order food check the weather listen to the radio set alarms and timers all of which is easily ac
complished through simple voice commands to me this is the exceedingly rare product that i didnt know i needed and now cant
live without similar to the smartphone and tablet i use every day that didnt exist just a few short years ago', 'if youre ev
en a little bit curious you owe it to yourself to give the dot a try add a good speaker and enjoy just how simple a connecte
d life can be', 'update after a bit more time with the dot or maybe i should say dots since i went out and bought another on
e for my living room ive come up with a few tips', ' use the best speakers you can with it i found that while bluetooth was
convenient i got much better sound out of my jbl duet computer speakers', ' take the time to voice train alexa at least once
its kinda tedious but really improves the accuracy ive now gone through three trainings with each dot the phrasing gets mo
re intricate with each and it really is amazing how much of an improvement it makes kinda hard to quantify but id guess alex
a is at least twice as likely to understand long complex phrases and has also gained noticeable accuracy when ambient noise i
```

We will remove all the garbage words from the list of reviews

We will remove all the garbage words from the list of reviews

```
In [13]: final_reviews = []
         for item in reviews:
             y = ''.join(x for x in item if ord(x) < 128)
             final_reviews.append(y)
```

```
In [14]: print(final_reviews)
```

```
['my brother robert who has been bed ridden and paralyzed with multiple sclerosis from his neck down for more than years no
w has a new friend named alexa he was in tears with happiness when alexa played s music played jeopardy answered all his que
stions and wakes him up every morning thank you amazon for giving my brother a new bedside companion', 'happy holidays', 'ro
y', 'having worked in the electronics retail industry for years now ive seen scores of smart devices come and go until now n
obody quite got it right in the echo dot amazon has created a near perfect blend of hardware and software ive seen plenty o
f the former but truly seamless multi platform software has eluded everyone but amazon were talking major players like samsu
ng and google who have been at it for much longer than amazon the main problem is that excellent products like the samsung s
mart things hub which do a fantastic job of unifying a slew of different connected devices from different companies nest hone
ywell phillips and so on still lacked the web connectivity and entertainment support i wanted so id still end up needing my t
ablet or phone thanks to fantastic third party support the dot has no problem controlling all of my smart stuff while allowi
ng me to listen to music order food check the weather listen to the radio set alarms and timers all of which is easily accomp
lished through simple voice commands to me this is the exceedingly rare product that i didnt know i needed and now cant live
without similar to the smartphone and tablet i use every day that didnt exist just a few short years ago', 'if youre even a
little bit curious you owe it to yourself to give the dot a try add a good speaker and enjoy just how simple a connected lif
e can be', 'update after a bit more time with the dot or maybe i should say dots since i went out and bought another one for
my living room ive come up with a few tips', ' use the best speakers you can with it i found that while bluetooth was conve
nient i got much better sound out of my jbl duet computer speakers', ' take the time to voice train alexa at least once its
kinda tedious but really improves the accuracy ive now gone through three trainings with each dot the phrasing gets more int
ricate with each and it really is amazing how much of an improvement it makes kinda hard to quantify but id guess alexa is a
t least twice as likely to understand long complex phrases and has also gained noticeable accuracy when ambient noise id once
```

We will strip all the white space from the list of reviews

We will strip all the white space from the list of reviews

```
In [15]: final_review = []  
        for item in reviews:  
            result = ' '.join(item.split())  
            final_review.append(result)
```

```
In [16]: print(final_review)
```

```
['my brother robert who has been bed ridden and paralyzed with multiple sclerosis from his neck down for more than years now  
has a new friend named alexa he was in tears with happiness when alexa played s music played jeopardy answered all his questi  
ons and wakes him up every morning thank you amazon for giving my brother a new bedside companion', 'happy holidays', 'roy',  
'having worked in the electronics retail industry for years now ive seen scores of smart devices come and go until now nobody  
quite got it right in the echo dot amazon has created a near perfect blend of hardware and software ive seen plenty of the fo  
rmer but truly seamless multi platform software has eluded everyone but amazon were talking major players like samsung and go  
ogle who have been at it for much longer than amazon the main problem is that excellent products like the samsung smart thing  
s hub which do a fantastic job of unifying a slew of different connected devices from different companies nest honeywell phil  
lips and so on still lacked the web connectivity and entertainment support i wanted so id still end up needing my tablet or p  
hone thanks to fantastic third party support the dot has no problem controlling all of my smart stuff while allowing me to li  
sten to music order food check the weather listen to the radio set alarms and timers all of which is easily accomplished thro  
ugh simple voice commands to me this is the exceedingly rare product that i didnt know i needed and now cant live without sim  
ilar to the smartphone and tablet i use every day that didnt exist just a few short years ago', 'if youre even a little bit c  
urious you owe it to yourself to give the dot a try add a good speaker and enjoy just how simple a connected life can be', 'u  
pdate after a bit more time with the dot or maybe i should say dots since i went out and bought another one for my living roo  
m ive come up with a few tips', 'use the best speakers you can with it i found that while bluetooth was convenient i got much  
better sound out of my jbl duet computer speakers', 'take the time to voice train alexa at least once its kinda tedious but r  
eally improves the accuracy ive now gone through three trainings with each dot the phrasing gets more intricate with each and  
it really is amazing how much of an improvement it makes kinda hard to quantify but id guess alexa is at least twice as likel  
y to understand long complex phrases and has also gained noticeable accuracy when ambient noise is present! 'if a phrase doe
```

We will import the NLTK library which we will use to remove the stop words from the list of reviews

```
In [17]: from nltk.corpus import stopwords
from nltk.tokenize import word_tokenize

stop_words = set(stopwords.words('english'))
print(stop_words)
```

```
{'didn', 'shouldn', 'had', 'weren', 'do', 'about', 'herself', 'o', 'with', 'or', 'i', 's', 'she', 'nor', 'shan', 'doing', 'yourself', 'which', 'aren', 'my', 'is', 'if', 'hers', 'and', 'was', 'were', 've', 'for', 'who', 'from', 'other', 're', 'be', 'these', 'does', 'just', 'than', 'your', 'has', 'wouldn', 'we', 'so', 'until', 'under', 'he', 'they', 'where', 'd', 'me', 'while', 'ma', 'did', 'it', 'her', 'being', 'by', 'll', 'their', 'what', 'all', 'too', 'm', 'hasn', 'mightn', 'up', 'that', 'whom', 'a', 'can', 'couldn', 'more', 'yours', 'both', 'you', 'out', 'but', 'most', 'hadn', 'when', 'above', 'mustn', 'only', 'same', 'itself', 'again', 'once', 'yourselves', 'ourselves', 'haven', 'the', 'few', 'own', 'any', 'him', 'as', 'against', 'doesn', 'of', 'so me', 'now', 'are', 'after', 'on', 'such', 'not', 'won', 'during', 'in', 'then', 'very', 'am', 'our', 'there', 'over', 'themselves', 'wasn', 'myself', 'each', 'himself', 'before', 'will', 'this', 'how', 'its', 'his', 'have', 'between', 'should', 'below', 'needn', 'an', 'why', 'having', 'to', 'at', 't', 'down', 'don', 'y', 'because', 'into', 'further', 'isn', 'been', 'ours', 'the m', 'theirs', 'here', 'those', 'off', 'ain', 'through', 'no'}
```

```
In [18]: amazon_reviews = []
for item in final_review:
    x = ' '.join([word for word in item.split() if word not in (stopwords.words('english'))])
    amazon_reviews.append(x)
```

```
In [19]: print(amazon_reviews)
```

```
['brother robert bed ridden paralyzed multiple sclerosis neck years new friend named alexa tears happiness alexa played music played jeopardy answered questions wakes every morning thank amazon giving brother new bedside companion', 'happy holidays', 'roy', 'worked electronics retail industry years ive seen scores smart devices come go nobody quite got right echo dot amazon created near perfect blend hardware software ive seen plenty former truly seamless multi platform software eluded everyone amazon talking major players like samsung google much longer amazon main problem excellent products like samsung smart things hub fantastic job unifying slew different connected devices different companies nest honeywell phillips still lacked web connectivity entertainment support wanted id still end needing tablet phone thanks fantastic third party support dot problem controlling smart stuff allowing listen music order food check weather listen radio set alarms timers easily accomplished simple voice commands exceedingly rare product didnt know needed cant live without similar smartphone tablet use every day didnt exist short years ago', 'youre even little bit curious owe give dot try add good speaker enjoy simple connected life', 'update bit time dot maybe say dots since went bought another one living room ive come tips', 'use best speakers found bluetooth convenient got much better sound jbl duet computer speakers', 'take time voice train alexa least kinda tedious really improves accuracy ive gone three trainings dot phrasing gets intricate really amazing much improvement makes kinda hard quantify id guess alexa least twice likely understand long complex phrases also gained noticeable accuracy ambient noise id present', 'phrase doesnt yield results youre looking reword try instance alexa lower temperature degrees got result tried alexa honeywell thermostat degrees picked perfectly', 'take time look skills theres lot helpful plain fun stuff strange facts calculator everything really helps enhance experience', 'id never really used prime music prior setting dots cant live without say basically whatever']
```

We will stem all the words so as to match it will the lexicons

We will stem all the words so as to match it will the lexicons

```
In [20]: from nltk.stem import PorterStemmer
from nltk.tokenize import sent_tokenize, word_tokenize

ps = PorterStemmer()

amazon_review = []
for item in amazon_reviews:
    x = ps.stem(item)
    amazon_review.append(x)
```

```
In [21]: print(amazon_review)
```

```
['brother robert bed ridden paralyzed multiple sclerosis neck years new friend named alexa tears happiness alexa played music played jeopardy answered questions wakes every morning thank amazon giving brother new bedside companion', 'happy holiday', 'roy', 'worked electronics retail industry years ive seen scores smart devices come go nobody quite got right echo dot amazon created near perfect blend hardware software ive seen plenty former truly seamless multi platform software eluded everyone amazon talking major players like samsung google much longer amazon main problem excellent products like samsung smart things hub fantastic job unifying slew different connected devices different companies nest honeywell phillips still lacked web connectivity entertainment support wanted id still end needing tablet phone thanks fantastic third party support dot problem controlling smart stuff allowing listen music order food check weather listen radio set alarms timers easily accomplished simple voice commands exceedingly rare product didnt know needed cant live without similar smartphone tablet use every day didnt exist short years ago', 'youre even little bit curious owe give dot try add good speaker enjoy simple connected lif', 'update bit time dot maybe say dots since went bought another one living room ive come tip', 'use best speakers found bluetooth convenient got much better sound jbl duet computer speak', 'take time voice train alexa least kinda tedious really improves accuracy ive gone three trainings dot phrasing gets intricate really amazing much improvement makes kinda hard quantify id guess alexa least twice likely understand long complex phrases also gained noticeable accuracy ambient noise id pres', 'phrase doesnt yield results youre looking reword try instance alexa lower temperature degrees got result tried alexa honeywell thermostat degrees picked perfectli', 'take time look skills theres lot helpful plain fun stuff strange facts calculator everything really helps enhance experi', 'id never really used prime music prior setting dots cant live without say basically whatever want get result favs far alexa play music alexa play indie music alexa play thunderstorm sounds last one ask repeat plays night really great freebie youre prime memb', 'bit worried initially alexa might triggered accidentally ambient tv general household noise im really impressed happened twice far times living room watching tv high volume concern mic temporarily disabled dot went to']
```



We will declare two empty list and then feed in the positive and negative words from the lexicons into the respective list

We will declare two empty list and then feed in the positive and negative words from the lexicons into the respective list

```
In [22]: neg_lexi = []
pos_lexi = []

with open("/Users/ninad/Desktop/ADS/Assignment08/positive-word.txt", errors='replace') as f:
    pos_lexi = f.readlines()
    pos_lexi = [x.strip() for x in pos_lexi]

with open("/Users/ninad/Desktop/ADS/Assignment08/negative-word.txt", errors='replace') as f:
    neg_lexi = f.readlines()
    neg_lexi = [x.strip() for x in neg_lexi]
```

We will import textblob to get words out of the sentences. Each review is then separated into words which is then passed on through the getSentiment function to get the sentiment either Positive or Negative or Neutral

We will import textblob to get words out of the sentences. Each review is then separated into words which is then passed on through the getSentiment function to get the sentiment either Positive or Negative or Neutral

```
In [23]: from textblob import TextBlob
import pandas as pd
```

```
In [24]: def getSentiment(s) :
    score = 0
    s_tb = TextBlob(s)
    s_tb = s_tb.words
    #filtered_s_tb = [w.upper() for w in s_tb if w.lower() not in stop_words]
    for i in s_tb:
        if i in pos_lexi:
            score += 1
        if i in neg_lexi:
            score -= 1
    if (score < 0) :
        response = "Negative"
    elif (score > 0) :
        response = "Positive"
    else :
        response = "Neutral"
    return response
```

```
In [25]: df = pd.DataFrame(columns = ['AmazonReview', 'Response'])
```

```
In [26]: for item in amazon_review:
    x = getSentiment(item)
    df = df.append(pd.DataFrame({'AmazonReview' : item, 'Response' : getSentiment(item)}, index=[0]), ignore_index=True)
```

```
In [27]: df.head()
```

```
Out[27]:
```

	AmazonReview	Response
0	brother robert bed ridden paralyzed multiple s...	Neutral
1	happy holiday	Positive
2	roy	Neutral
3	worked electronics retail industrv years ive s	Positive

```
In [28]: df_neg = df[df['Response'] == 'Negative']
```

```
In [29]: df_neg.head()
```

```
Out[29]:
```

	AmazonReview	Response
11	bit worried initially alexa might triggered ac...	Negative
19	dont normally write reviews time recently lost...	Negative
38	whether want weather latest sports scores help...	Negative
49	smarthome devices named internal functions get...	Negative
53	main selling point device use exsisting blueto...	Negative

```
In [30]: df.to_csv('LexiconResult.csv', sep=',')
```

To pass it through an NBC, we will need to convert the dataframe into a list. An item in the list will have its review as item[0] and the sentiment as item[1]

We will now divide this labelled list into training and testing sets, with training having 4000 reviews and testing having 1000 reviews

We will import the Multinomial NBC from the sklearn module

## NBC

To pass it through an NBC, we will need to convert the dataframe into a list. An item in the list will have its review as item[0] and the sentiment as item[1]

```
In [31]: l1 = df.values.tolist()
```

We will now divide this labelled list into training and testing sets, with training having 4000 reviews and testing having 1000 reviews

```
In [42]: training_set = l1[:3000]
testing_set = l1[1000:]
```

We will import the Multinomial NBC from the sklearn module

```
In [43]: from sklearn.naive_bayes import MultinomialNB
from sklearn.feature_extraction.text import CountVectorizer
from sklearn import metrics
```



Generate counts from text using a vectorizer. We can choose from other available vectorizers, and set many different options. This code performs our step of computing word counts

## NBC

To pass it through an NBC, we will need to convert the dataframe into a list. An item in the list will have its review as item[0] and the sentiment as item[1]

```
In [32]: l1 = df.values.tolist()
```

✖ Rectangular Snip

We will now divide this labelled list into training and testing sets, with training having 4000 reviews and testing having 1000 reviews

```
In [33]: training_set = l1[:3000]
testing_set = l1[3000:]
```

We will import the Multinomial NBC from the sklearn module. An easier way to use Naive Bayes is to use the implementation in scikit-learn. Scikit-learn is a Python machine learning library that contains implementations of all the common machine learning algorithms

```
In [34]: from sklearn.naive_bayes import MultinomialNB
from sklearn.feature_extraction.text import CountVectorizer
from sklearn import metrics
```

```
In [35]: # Generate counts from text using a vectorizer
# We can choose from other available vectorizers, and set many different options
# This code performs our step of computing word counts
vectorizer = CountVectorizer(stop_words='english', max_df=.05)
train_features = vectorizer.fit_transform([r[0] for r in training_set])
test_features = vectorizer.transform([r[0] for r in testing_set])
```

```
In [36]: # Fit a Naive Bayes model to the training data
# This will train the model using the word counts we computed and the existing classifications in the training set
nb = MultinomialNB()
nb.fit(train_features, [r[1] for r in training_set])
```

```
Out[36]: MultinomialNB(alpha=1.0, class_prior=None, fit_prior=True)
```

```
In [37]: # Now we can use the model to predict classifications for our test features
predictions = nb.predict(test_features)
df_predict = pd.DataFrame(columns = ['Result_Predicted'])
for item in predictions:
    df_predict = df_predict.append(pd.DataFrame({'Result_Predicted' : item}, index=[0]), ignore_index=True)
```

```
In [38]: df_new = pd.DataFrame(columns = ['Predicted', 'Result_Actual'])
for item in testing_set:
    #x = cl.classify(item)
    df_new= df_new.append(pd.DataFrame({'Predicted' : item[0], 'Result_Actual' : item[1]}, index=[0]), ignore_index=True)
```

```
In [39]: df_result = pd.concat([df_new, df_predict], axis=1)
```

```
In [40]: df_result.head()
```

```
Out[40]:
```

	Predicted	Result_Actual	Result_Predicted
0	fell deep like echo two included security syst...	Positive	Positive
1	little gadget awesome gave stars says works we...	Positive	Positive
2	dont know went wrong fulsized echo great resp...	Neutral	Positive
3	thing probably worst amazon product ive ever p...	Negative	Positive
4	edit i'm changing review swapped trigger word ...	Neutral	Neutral

```
In [41]: df_result.to_csv('NBCResult.csv', sep=',')
```

---

Calculate the accuracy of NBC

```
In [42]: rows = list(df_result[df_result['Result_Actual'] == df_result['Result_Predicted']].index)
len(rows)
```

```
Out[42]: 7135
```

```
In [43]: len(df_result.index)
```

```
Out[43]: 10002
```

```
In [44]: print(len(rows)/len(df_result.index) *100)
71.3357328534293
```

The Accuracy I get from NB is **71.33%**