runfile('C:/Users/kryp\_/Documents/conrad/MSDSC/DSC680/proj-TwitterSentiment/src/main.py', wdir='C:/Users/kryp\_/Documents/conrad/MSDSC/DSC680/proj-TwitterSentiment/src')

109924

In graph\_data

train\_data textID ... sentiment

0 cb774db0d1 ... neutral

1 549e992a42 ... negative

2 088c60f138 ... negative

3 9642c003ef ... negative

4 358bd9e861 ... negative

[5 rows x 4 columns]

positive\_data textID ... sentiment

6 6e0c6d75b1 ... positive

9 fc2cbefa9d ... positive

11 16fab9f95b ... positive

21 e48b0b8a23 ... positive

25 e00c6ef376 ... positive

[5 rows x 4 columns]

negative\_data textID ... sentiment

1 549e992a42 ... negative

2 088c60f138 ... negative

3 9642c003ef ... negative

4 358bd9e861 ... negative

12 74a76f6e0a ... negative

[5 rows x 4 columns]

neutral\_data textID ... sentiment

0 cb774db0d1 ... neutral

5 28b57f3990 ... neutral

7 50e14c0bb8 ... neutral

8 e050245fbd ... neutral

10 2339a9b08b ... neutral

[5 rows x 4 columns]

Positive

[('good', 950), ('love', 795), ('happy', 733), ('day', 476), ('thanks', 438), ('great', 370), ('fun', 287), ('nice', 268), ('hope', 247), ('awesome', 234)]

Negative

[('miss', 525), ('sad', 347), ('sorry', 294), ('bad', 294), ('..', 265), ('hate', 248), ('...', 230), ('feel', 219), ('get', 219), ('suck', 208)]

Neutral

[('...', 1392), ('get', 1239), ('go', 1147), ('..', 638), ('day', 629), ('work', 625), ('like', 462), ('lol', 462), ('time', 458), ('know', 442)]

Train size: 5496

Training data: ({'wish': True}, 'positive')

Training data: ({'man': True, 'great': True, 'sense': True, 'humour': True, '..': True}, 'positive')

Test data: ({'cute': True}, 'positive')

Test data: ({'tire': True}, 'negative')

Accuracy is: 0.7075733454628156

Most Informative Features

love = True positi : negati = 59.6 : 1.0

miss = True negati : positi = 37.0 : 1.0

hurt = True negati : positi = 35.8 : 1.0

http = True neutra : negati = 31.8 : 1.0

happy = True positi : negati = 30.8 : 1.0

best = True positi : negati = 19.7 : 1.0

hey = True neutra : positi = 18.4 : 1.0

thanks = True positi : negati = 18.0 : 1.0

tonight = True neutra : negati = 17.2 : 1.0

watch = True neutra : negati = 16.8 : 1.0

hour = True neutra : positi = 16.8 : 1.0

next = True neutra : negati = 16.7 : 1.0

yes = True neutra : negati = 16.3 : 1.0

call = True neutra : positi = 16.3 : 1.0

find = True neutra : positi = 16.3 : 1.0

play = True neutra : positi = 16.3 : 1.0

sorry = True negati : positi = 15.7 : 1.0

e = True positi : neutra = 15.4 : 1.0

sick = True negati : positi = 15.2 : 1.0

twitter = True neutra : positi = 14.9 : 1.0

None

Evaluating Actual Test Data File

10602

In graph\_data

Traceback (most recent call last):

File "C:\Users\kryp\_\Documents\conrad\MSDSC\DSC680\proj-TwitterSentiment\src\main.py", line 324, in <module>

main()

File "C:\Users\kryp\_\Documents\conrad\MSDSC\DSC680\proj-TwitterSentiment\src\main.py", line 230, in main

create\_word\_cloud(actual\_positive\_cleaned\_tokens\_list, stop\_word, "TestPositiveWordCloud")

File "C:\Users\kryp\_\Documents\conrad\MSDSC\DSC680\proj-TwitterSentiment\src\main.py", line 47, in create\_word\_cloud

wordcloud = WordCloud(stopwords=stopwords).generate(full\_text)

File "C:\Anaconda\lib\site-packages\wordcloud\wordcloud.py", line 632, in generate

return self.generate\_from\_text(text)

File "C:\Anaconda\lib\site-packages\wordcloud\wordcloud.py", line 614, in generate\_from\_text

self.generate\_from\_frequencies(words)

File "C:\Anaconda\lib\site-packages\wordcloud\wordcloud.py", line 541, in generate\_from\_frequencies

occupancy.update(img\_array, x, y)

File "C:\Anaconda\lib\site-packages\wordcloud\wordcloud.py", line 55, in update

axis=1), axis=0)

File "<\_\_array\_function\_\_ internals>", line 6, in cumsum

File "C:\Anaconda\lib\site-packages\numpy\core\fromnumeric.py", line 2483, in cumsum

return \_wrapfunc(a, 'cumsum', axis=axis, dtype=dtype, out=out)

File "C:\Anaconda\lib\site-packages\numpy\core\fromnumeric.py", line 58, in \_wrapfunc

return bound(\*args, \*\*kwds)

KeyboardInterrupt

runfile('C:/Users/kryp\_/Documents/conrad/MSDSC/DSC680/proj-TwitterSentiment/src/main.py', wdir='C:/Users/kryp\_/Documents/conrad/MSDSC/DSC680/proj-TwitterSentiment/src')

109924

In graph\_data

train\_data textID ... sentiment

0 cb774db0d1 ... neutral

1 549e992a42 ... negative

2 088c60f138 ... negative

3 9642c003ef ... negative

4 358bd9e861 ... negative

[5 rows x 4 columns]

positive\_data textID ... sentiment

6 6e0c6d75b1 ... positive

9 fc2cbefa9d ... positive

11 16fab9f95b ... positive

21 e48b0b8a23 ... positive

25 e00c6ef376 ... positive

[5 rows x 4 columns]

negative\_data textID ... sentiment

1 549e992a42 ... negative

2 088c60f138 ... negative

3 9642c003ef ... negative

4 358bd9e861 ... negative

12 74a76f6e0a ... negative

[5 rows x 4 columns]

neutral\_data textID ... sentiment

0 cb774db0d1 ... neutral

5 28b57f3990 ... neutral

7 50e14c0bb8 ... neutral

8 e050245fbd ... neutral

10 2339a9b08b ... neutral

[5 rows x 4 columns]

Positive

[('good', 950), ('love', 795), ('happy', 733), ('day', 476), ('thanks', 438), ('great', 370), ('fun', 287), ('nice', 268), ('hope', 247), ('awesome', 234)]

Negative

[('miss', 525), ('sad', 347), ('sorry', 294), ('bad', 294), ('..', 265), ('hate', 248), ('...', 230), ('feel', 219), ('get', 219), ('suck', 208)]

Neutral

[('...', 1392), ('get', 1239), ('go', 1147), ('..', 638), ('day', 629), ('work', 625), ('like', 462), ('lol', 462), ('time', 458), ('know', 442)]

Train size: 5496

Training data: ({'wish': True}, 'positive')

Training data: ({'man': True, 'great': True, 'sense': True, 'humour': True, '..': True}, 'positive')

Test data: ({'cute': True}, 'positive')

Test data: ({'tire': True}, 'negative')

Accuracy is: 0.7075733454628156

Most Informative Features

love = True positi : negati = 59.6 : 1.0

miss = True negati : positi = 37.0 : 1.0

hurt = True negati : positi = 35.8 : 1.0

http = True neutra : negati = 31.8 : 1.0

happy = True positi : negati = 30.8 : 1.0

best = True positi : negati = 19.7 : 1.0

hey = True neutra : positi = 18.4 : 1.0

thanks = True positi : negati = 18.0 : 1.0

tonight = True neutra : negati = 17.2 : 1.0

watch = True neutra : negati = 16.8 : 1.0

hour = True neutra : positi = 16.8 : 1.0

next = True neutra : negati = 16.7 : 1.0

yes = True neutra : negati = 16.3 : 1.0

call = True neutra : positi = 16.3 : 1.0

find = True neutra : positi = 16.3 : 1.0

play = True neutra : positi = 16.3 : 1.0

sorry = True negati : positi = 15.7 : 1.0

e = True positi : neutra = 15.4 : 1.0

sick = True negati : positi = 15.2 : 1.0

twitter = True neutra : positi = 14.9 : 1.0

None

Evaluating Actual Test Data File

10602

In graph\_data

Traceback (most recent call last):

File "C:\Users\kryp\_\Documents\conrad\MSDSC\DSC680\proj-TwitterSentiment\src\main.py", line 324, in <module>

main()

File "C:\Users\kryp\_\Documents\conrad\MSDSC\DSC680\proj-TwitterSentiment\src\main.py", line 230, in main

create\_word\_cloud(actual\_positive\_cleaned\_tokens\_list, stop\_word, "TestPositiveWordCloud")

File "C:\Users\kryp\_\Documents\conrad\MSDSC\DSC680\proj-TwitterSentiment\src\main.py", line 47, in create\_word\_cloud

wordcloud = WordCloud(stopwords=stopwords).generate(full\_text)

File "C:\Anaconda\lib\site-packages\wordcloud\wordcloud.py", line 632, in generate

return self.generate\_from\_text(text)

File "C:\Anaconda\lib\site-packages\wordcloud\wordcloud.py", line 614, in generate\_from\_text

self.generate\_from\_frequencies(words)

File "C:\Anaconda\lib\site-packages\wordcloud\wordcloud.py", line 541, in generate\_from\_frequencies

occupancy.update(img\_array, x, y)

File "C:\Anaconda\lib\site-packages\wordcloud\wordcloud.py", line 55, in update

axis=1), axis=0)

File "<\_\_array\_function\_\_ internals>", line 6, in cumsum

File "C:\Anaconda\lib\site-packages\numpy\core\fromnumeric.py", line 2483, in cumsum

return \_wrapfunc(a, 'cumsum', axis=axis, dtype=dtype, out=out)

File "C:\Anaconda\lib\site-packages\numpy\core\fromnumeric.py", line 58, in \_wrapfunc

return bound(\*args, \*\*kwds)

KeyboardInterrupt

runfile('C:/Users/kryp\_/Documents/conrad/MSDSC/DSC680/proj-TwitterSentiment/src/main.py', wdir='C:/Users/kryp\_/Documents/conrad/MSDSC/DSC680/proj-TwitterSentiment/src')

109924

In graph\_data

train\_data textID ... sentiment

0 cb774db0d1 ... neutral

1 549e992a42 ... negative

2 088c60f138 ... negative

3 9642c003ef ... negative

4 358bd9e861 ... negative

[5 rows x 4 columns]

positive\_data textID ... sentiment

6 6e0c6d75b1 ... positive

9 fc2cbefa9d ... positive

11 16fab9f95b ... positive

21 e48b0b8a23 ... positive

25 e00c6ef376 ... positive

[5 rows x 4 columns]

negative\_data textID ... sentiment

1 549e992a42 ... negative

2 088c60f138 ... negative

3 9642c003ef ... negative

4 358bd9e861 ... negative

12 74a76f6e0a ... negative

[5 rows x 4 columns]

neutral\_data textID ... sentiment

0 cb774db0d1 ... neutral

5 28b57f3990 ... neutral

7 50e14c0bb8 ... neutral

8 e050245fbd ... neutral

10 2339a9b08b ... neutral

[5 rows x 4 columns]

Positive

[('good', 950), ('love', 795), ('happy', 733), ('day', 476), ('thanks', 438), ('great', 370), ('fun', 287), ('nice', 268), ('hope', 247), ('awesome', 234)]

Negative

[('miss', 525), ('sad', 347), ('sorry', 294), ('bad', 294), ('..', 265), ('hate', 248), ('...', 230), ('feel', 219), ('get', 219), ('suck', 208)]

Neutral

[('...', 1392), ('get', 1239), ('go', 1147), ('..', 638), ('day', 629), ('work', 625), ('like', 462), ('lol', 462), ('time', 458), ('know', 442)]

Train size: 5496

Training data: ({'wish': True}, 'positive')

Training data: ({'man': True, 'great': True, 'sense': True, 'humour': True, '..': True}, 'positive')

Test data: ({'cute': True}, 'positive')

Test data: ({'tire': True}, 'negative')

Accuracy is: 0.7075733454628156

Most Informative Features

love = True positi : negati = 59.6 : 1.0

miss = True negati : positi = 37.0 : 1.0

hurt = True negati : positi = 35.8 : 1.0

http = True neutra : negati = 31.8 : 1.0

happy = True positi : negati = 30.8 : 1.0

best = True positi : negati = 19.7 : 1.0

hey = True neutra : positi = 18.4 : 1.0

thanks = True positi : negati = 18.0 : 1.0

tonight = True neutra : negati = 17.2 : 1.0

watch = True neutra : negati = 16.8 : 1.0

hour = True neutra : positi = 16.8 : 1.0

next = True neutra : negati = 16.7 : 1.0

yes = True neutra : negati = 16.3 : 1.0

call = True neutra : positi = 16.3 : 1.0

find = True neutra : positi = 16.3 : 1.0

play = True neutra : positi = 16.3 : 1.0

sorry = True negati : positi = 15.7 : 1.0

e = True positi : neutra = 15.4 : 1.0

sick = True negati : positi = 15.2 : 1.0

twitter = True neutra : positi = 14.9 : 1.0

None

Evaluating Actual Test Data File

10602

In graph\_data

Accuracy is: 0.7847715228477152

Most Informative Features

sad = True negati : positi = 151.5 : 1.0

thank = True positi : negati = 109.1 : 1.0

hate = True negati : positi = 105.2 : 1.0

suck = True negati : positi = 92.0 : 1.0

awesome = True positi : negati = 84.0 : 1.0

http = True neutra : positi = 76.0 : 1.0

stupid = True negati : positi = 65.8 : 1.0

fail = True negati : positi = 51.1 : 1.0

beautiful = True positi : negati = 46.2 : 1.0

till = True neutra : positi = 44.0 : 1.0

amazing = True positi : negati = 42.6 : 1.0

thanks = True positi : negati = 41.7 : 1.0

sick = True negati : positi = 40.8 : 1.0

love = True positi : negati = 40.5 : 1.0

hour = True neutra : positi = 35.4 : 1.0

hurt = True negati : positi = 34.6 : 1.0

poor = True negati : positi = 33.7 : 1.0

close = True neutra : positi = 33.2 : 1.0

lose = True negati : positi = 32.9 : 1.0

mother = True positi : negati = 32.6 : 1.0

None

Making adjustments to remove …

runfile('C:/Users/kryp\_/Documents/conrad/MSDSC/DSC680/proj-TwitterSentiment/src/main.py', wdir='C:/Users/kryp\_/Documents/conrad/MSDSC/DSC680/proj-TwitterSentiment/src')

109924

In graph\_data

train\_data textID ... sentiment

0 cb774db0d1 ... neutral

1 549e992a42 ... negative

2 088c60f138 ... negative

3 9642c003ef ... negative

4 358bd9e861 ... negative

[5 rows x 4 columns]

positive\_data textID ... sentiment

6 6e0c6d75b1 ... positive

9 fc2cbefa9d ... positive

11 16fab9f95b ... positive

21 e48b0b8a23 ... positive

25 e00c6ef376 ... positive

[5 rows x 4 columns]

negative\_data textID ... sentiment

1 549e992a42 ... negative

2 088c60f138 ... negative

3 9642c003ef ... negative

4 358bd9e861 ... negative

12 74a76f6e0a ... negative

[5 rows x 4 columns]

neutral\_data textID ... sentiment

0 cb774db0d1 ... neutral

5 28b57f3990 ... neutral

7 50e14c0bb8 ... neutral

8 e050245fbd ... neutral

10 2339a9b08b ... neutral

[5 rows x 4 columns]

Positive

[('i', 1215), ('good', 950), ('be', 860), ('love', 795), ('happy', 733), ('you', 705), ('to', 608), ('a', 590), ('the', 571), ('s', 482)]

Negative

[('i', 1713), ('be', 887), ('t', 596), ('to', 595), ('the', 547), ('my', 527), ('miss', 525), ('a', 473), ('it', 431), ('not', 420)]

Neutral

[('i', 6078), ('to', 4104), ('be', 3921), ('the', 3480), ('a', 2481), ('my', 1979), ('it', 1974), ('you', 1918), ('and', 1832), ('in', 1573)]

Train size: 5496

Training data: ({'wish': True}, 'positive')

Training data: ({'man': True, 'with': True, 'a': True, 'great': True, 'sense': True, 'of': True, 'humour': True}, 'positive')

Test data: ({'cute': True}, 'positive')

Test data: ({'tire': True}, 'negative')

Accuracy is: 0.6556288378439845

Most Informative Features

love = True positi : negati = 59.6 : 1.0

miss = True negati : positi = 37.0 : 1.0

hurt = True negati : positi = 35.8 : 1.0

http = True neutra : negati = 31.8 : 1.0

happy = True positi : negati = 30.8 : 1.0

only = True neutra : negati = 24.7 : 1.0

over = True neutra : positi = 23.7 : 1.0

or = True neutra : positi = 23.2 : 1.0

if = True neutra : positi = 20.3 : 1.0

best = True positi : negati = 19.7 : 1.0

hey = True neutra : positi = 18.4 : 1.0

up = True neutra : positi = 18.2 : 1.0

thanks = True positi : negati = 18.0 : 1.0

some = True neutra : negati = 18.0 : 1.0

after = True neutra : positi = 17.9 : 1.0

tonight = True neutra : negati = 17.2 : 1.0

watch = True neutra : negati = 16.8 : 1.0

hour = True neutra : positi = 16.8 : 1.0

next = True neutra : negati = 16.7 : 1.0

yes = True neutra : negati = 16.3 : 1.0

None

Evaluating Actual Test Data File

10602

In graph\_data

Accuracy is: 0.7365763423657634

Most Informative Features

sad = True negati : positi = 151.5 : 1.0

thank = True positi : negati = 109.1 : 1.0

hate = True negati : positi = 105.2 : 1.0

suck = True negati : positi = 92.0 : 1.0

awesome = True positi : negati = 84.0 : 1.0

http = True neutra : positi = 76.0 : 1.0

stupid = True negati : positi = 65.8 : 1.0

before = True neutra : positi = 56.9 : 1.0

fail = True negati : positi = 51.1 : 1.0

beautiful = True positi : negati = 46.2 : 1.0

other = True neutra : positi = 46.1 : 1.0

till = True neutra : positi = 44.0 : 1.0

amazing = True positi : negati = 42.6 : 1.0

thanks = True positi : negati = 41.7 : 1.0

sick = True negati : positi = 40.8 : 1.0

love = True positi : negati = 40.5 : 1.0

hour = True neutra : positi = 35.4 : 1.0

hurt = True negati : positi = 34.6 : 1.0

poor = True negati : positi = 33.7 : 1.0

close = True neutra : positi = 33.2 : 1.0

None