Report: Natural Language Processing

Identifying Body Shaming Tweets

Body shaming is a big problem in social media. People, especially celebrities or social media influencers, body positivity models are subjected to harsh body shaming comments. This is an attempt to identify body shaming comments which might potentially help companies flag such comments to people.

Data

I am using the Twitter Sentiment Analysis Dataset, which can be found here. https://www.kaggle.com/datasets/kazanova/sentiment140. I am using this dataset to build a model that identifies body shaming tweets. I am using twitter data because it was readily available. The dataset contains 1 million tweets which could be used to develop NLP models.

Data Wrangling

I used the following steps to clean and wrangle data. The data already had a lot of tweets. I did not need to merge it with another dataset.

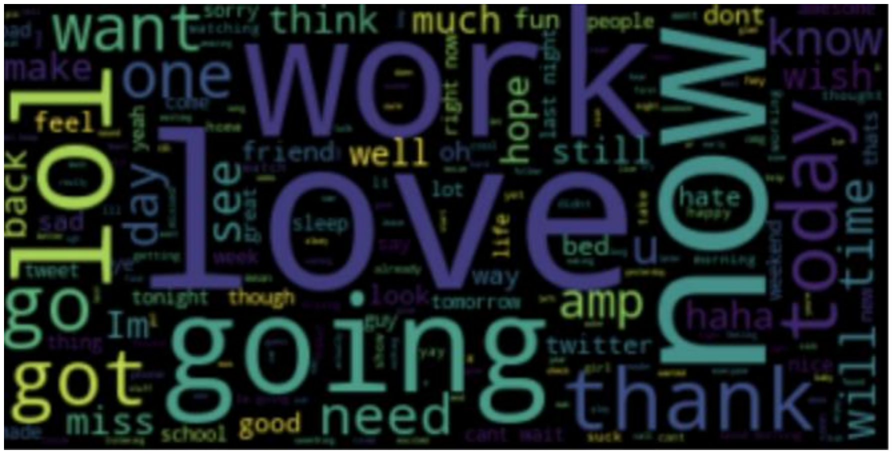
Step 1: To ensure that the texts are analyzable and readable, first, I changed the datatype to string.

Step 2: Then, I removed numbers or symbols from the text because they were of little analytical importance.

Exploratory Data Analysis

I took the following steps to explore the data.

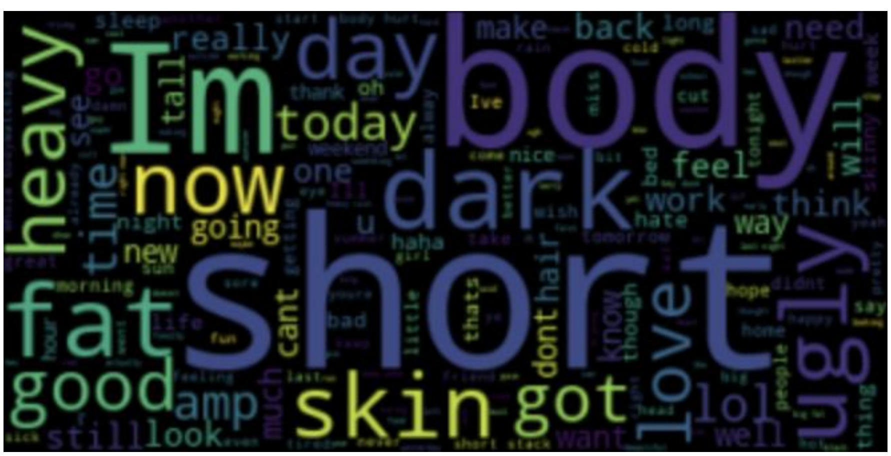
1. First, I explored the corpus of my texts to see which words were most common in the entire corpus. This also gave me an idea of how common tweets related to body shaming were in the corpus. I created the following word cloud.



The most common words here are not of interest to my analysis. I will therefore look at tweets that contain some of the most common body shaming related words and generate a word cloud and bar graph out of those tweets.

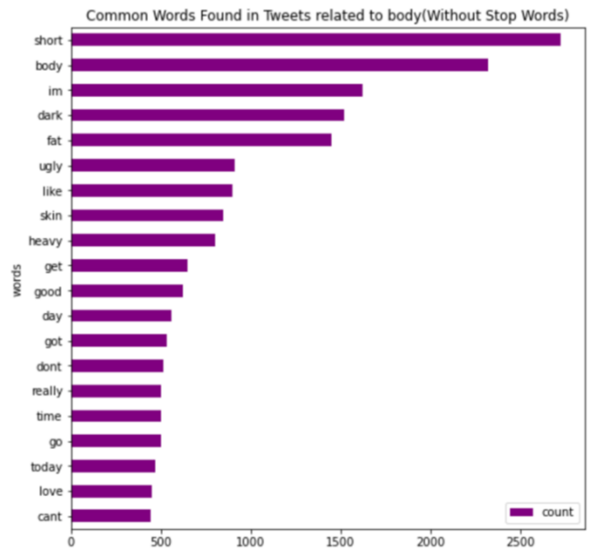
2. I filtered those tweets that contained the following general words that are related to body shaming contents. These words were identified based on my google search of body shaming tweets and from reading about body positivity influencers who are subjected to body shaming comments. I filtered tweets that contained: 'ugly', 'fat', 'body', 'dark', 'heavy', 'jiggle', 'short', 'tall', 'skinny', 'skin'. The filtered dataset contained 10782 tweets.

   I then generated the word cloud below.



Words such as ‘dark’, ‘body’, ‘ugly’, ‘short’, ‘fat’ are quite common in these tweets.

3. Next, I created a bar graph showing the most common words. ‘Short’, ‘body’, ‘dark’, ‘fat’, ‘ugly’ are very high in the list.



Pre-processing

I used the Count Vectorizer and TFIDF Vectorizer to vectorize the corpus.

1. **Count Vectorizer** showed that in our filtered dataset with some key words of interest, the word 'fat' appears 1452 times, 'ugly' 909 times, 'body' 2327 times, 'skin' 851 times, and 'dark' 1522 times.
2. I also used the count vectorizer with some bigrams and trigrams. It is interesting to see if certain phrases or combination of words were more common. These were some of the most common phrases/words.

Words/Phrases	Count
'ugly body'	32
'fat body'	11
'dark skin'	9
'hate body'	20
'heavy body'	22
'big tummy'	3
'fat chick'	13
'fat girl'	14
'fat boy'	12
'fat guy'	15
'short chick'	26
'short guy'	27
'short man'	28

Phrases	Count
You are ugly	21
You are fat	13
You are dark	12

You are jiggling	16
You are gross	14
‘fat chick gross’	14
You are hairy	15
Your heavy body	25
Your fat butt	24
Your fat ass	23
You are short	17
You are tall	20
You are short dude	18
You are short guy	19

3. I also used **TFIDF vectorizer** and compare it to the count vectorizer. The tfidf vectorizer will count the words and give importance to the words based on how many times it occurs throughout the corpus.

   ‘ugly’: 273, ‘fat’: 361, ‘body’: 488, skin: ‘228’, ‘dark’: 358

4. I also used bigrams and trigrams with the tfidf vectorizer.

Words/phrases	Weight
‘ugly body’	0.0
‘fat body’	1.48
‘dark skin’	0.78
‘big body’	0.69
‘hate body’	8.61
‘fat body’	8.10

5. The tfidf vectorizer did not give as much importance to trigrams.

Modeling

Running the LDA Model for Topic Identification

1. Next, I ran the LDA model for topic identification, to see whether these tweets can either be classified as 'body' or 'not body'. Then, combined with the sentiment, I determined whether these tweets are body shaming or not. If the sentiment is negative and is classified as 'body', the tweets are classified as 'body shaming'. If the sentiment is positive or neutral and classified as 'body', it is classified as 'not body shaming'.

   I used both the ‘Bag of Words’ and ‘Tfidf Model’ to identify topics.

‘Bag of Words’ Model:

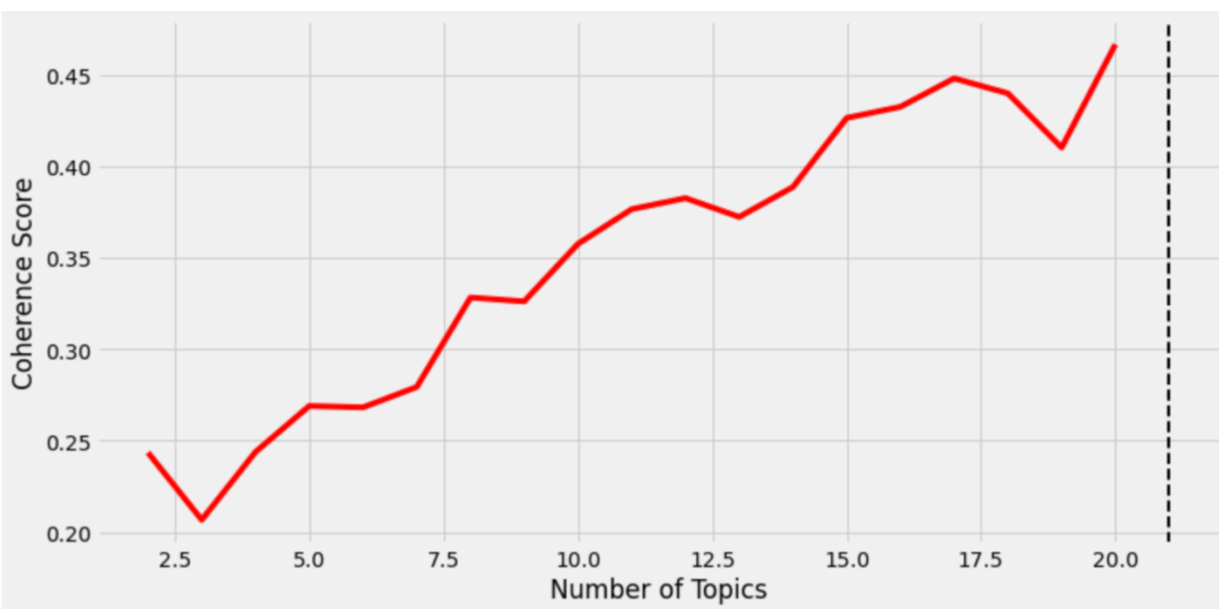
```
Topic: 0
Words: 0.037*"short" + 0.032*"skin" + 0.032*"bodi" + 0.028*"dark" + 0.021*"know" + 0.018*"like" + 0.017*"peopl" + 0.0
15*"burn" + 0.014*"feel" + 0.014*"go"
Topic: 1
Words: 0.055*"bodi" + 0.031*"tall" + 0.029*"need" + 0.024*"dark" + 0.022*"ugli" + 0.018*"like" + 0.017*"love" + 0.015
*"heavi" + 0.013*"say" + 0.012*"short"
Topic: 2
Words: 0.111*"dark" + 0.028*"bodi" + 0.023*"haha" + 0.021*"want" + 0.020*"good" + 0.018*"watch" + 0.016*"today" + 0.0
15*"night" + 0.014*"need" + 0.013*"dont"
Topic: 3
Words: 0.063*"heavi" + 0.059*"short" + 0.031*"work" + 0.029*"bodi" + 0.026*"week" + 0.024*"thank" + 0.022*"rain" + 0.
022*"ugli" + 0.020*"today" + 0.020*"feel"
Topic: 4
Words: 0.093*"short" + 0.060*"dark" + 0.025*"time" + 0.020*"today" + 0.019*"go" + 0.017*"think" + 0.015*"skin" + 0.01
5*"get" + 0.014*"bodi" + 0.010*"week"
Topic: 5
Words: 0.081*"short" + 0.038*"bodi" + 0.031*"heavi" + 0.019*"think" + 0.016*"skinni" + 0.015*"wait" + 0.014*"time" +
0.013*"stack" + 0.013*"work" + 0.012*"love"
Topic: 6
Words: 0.078*"bodi" + 0.063*"short" + 0.025*"like" + 0.023*"dont" + 0.021*"go" + 0.017*"hair" + 0.017*"look" + 0.016
*"skinni" + 0.015*"know" + 0.014*"good"
Topic: 7
Words: 0.072*"short" + 0.048*"love" + 0.040*"like" + 0.038*"ugli" + 0.028*"hair" + 0.027*"skin" + 0.023*"good" + 0.01
8*"dark" + 0.016*"go" + 0.015*"look"
Topic: 8
Words: 0.068*"short" + 0.033*"dark" + 0.019*"sleep" + 0.018*"work" + 0.017*"ugli" + 0.017*"bodi" + 0.015*"want" + 0.0
15*"awesom" + 0.014*"long" + 0.012*"night"
Topic: 9
Words: 0.106*"bodi" + 0.033*"skin" + 0.031*"think" + 0.029*"like" + 0.027*"feel" + 0.025*"short" + 0.019*"hurt" + 0.0
14*"dont" + 0.014*"life" + 0.014*"look"
```

‘Tfidf Model’

```
Topic: 0 Word: 0.044*"bodi" + 0.041*"short" + 0.023*"dark" + 0.020*"get" + 0.014*"time" + 0.012*"week" + 0.010*"nigh
t" + 0.010*"love" + 0.010*"have" + 0.009*"work"
Topic: 1 Word: 0.033*"short" + 0.025*"heavi" + 0.019*"like" + 0.015*"say" + 0.014*"bodi" + 0.014*"rain" + 0.011*"ski
n" + 0.009*"long" + 0.009*"need" + 0.009*"dark"
Topic: 2 Word: 0.044*"skin" + 0.017*"bodi" + 0.016*"good" + 0.015*"know" + 0.015*"readi" + 0.013*"dark" + 0.012*"shor
t" + 0.012*"tonight" + 0.011*"tri" + 0.011*"yeah"
Topic: 3 Word: 0.040*"ugli" + 0.029*"dark" + 0.021*"tall" + 0.019*"short" + 0.017*"time" + 0.015*"like" + 0.013*"bod
i" + 0.010*"think" + 0.010*"great" + 0.009*"meet"
Topic: 4 Word: 0.032*"dark" + 0.017*"short" + 0.017*"bodi" + 0.016*"know" + 0.014*"go" + 0.014*"right" + 0.013*"choco
l" + 0.013*"hair" + 0.011*"come" + 0.011*"sexi"
Topic: 5 Word: 0.031*"love" + 0.020*"bodi" + 0.018*"short" + 0.015*"tall" + 0.015*"good" + 0.013*"heavi" + 0.013*"lik
e" + 0.012*"ugli" + 0.010*"go" + 0.010*"dark"
Topic: 6 Word: 0.025*"short" + 0.018*"bodi" + 0.017*"hate" + 0.015*"dark" + 0.014*"weekend" + 0.013*"nice" + 0.013*"f
eel" + 0.013*"want" + 0.013*"like" + 0.012*"ugli"
Topic: 7 Word: 0.028*"skinni" + 0.021*"short" + 0.017*"bodi" + 0.014*"look" + 0.013*"dont" + 0.013*"dark" + 0.013*"g
o" + 0.013*"work" + 0.012*"haha" + 0.011*"jean"
Topic: 8 Word: 0.038*"short" + 0.035*"hair" + 0.016*"heavi" + 0.015*"today" + 0.013*"bodi" + 0.012*"love" + 0.012*"da
rk" + 0.011*"skin" + 0.011*"ugli" + 0.010*"coffe"
Topic: 9 Word: 0.019*"short" + 0.016*"good" + 0.015*"dark" + 0.015*"think" + 0.015*"bodi" + 0.015*"work" + 0.014*"don
t" + 0.012*"love" + 0.012*"that" + 0.012*"peopl"
```

The two models show that some topics are related to body and some not. Next, I ran the coherence scores to find the optimum number of topics.

	Number of Topics	Coherence Score
18	20	0.4667
15	17	0.4481
16	18	0.4399
14	16	0.4326
13	15	0.4265
17	19	0.4103
12	14	0.3887
10	12	0.3827
9	11	0.3766
11	13	0.3724



Although the figure shows that 20 is the most optimal number of topics, I am going with 8 topics because the figure shows the highest jump from previous topic to the next one is from 7.5 to 8. So, 8 topics carry the most analytical importance.

Re-running the LDA model with 8 topics, I get the following.

```
Topic: 0 Word: 0.037*"dark" + 0.025*"look" + 0.017*"short" + 0.014*"bodi" + 0.014*"skin" + 0.014*"go" + 0.013*"miss"
+ 0.012*"work" + 0.011*"hair" + 0.010*"say"
Topic: 1 Word: 0.024*"skin" + 0.020*"heavi" + 0.019*"feel" + 0.016*"dark" + 0.016*"skinni" + 0.015*"short" + 0.013*"r
ain" + 0.012*"jean" + 0.011*"bodi" + 0.011*"love"
Topic: 2 Word: 0.049*"short" + 0.023*"like" + 0.015*"love" + 0.014*"bodi" + 0.013*"dark" + 0.012*"watch" + 0.010*"wee
k" + 0.010*"think" + 0.010*"tall" + 0.009*"long"
Topic: 3 Word: 0.023*"bodi" + 0.022*"dark" + 0.013*"skin" + 0.013*"short" + 0.012*"hurt" + 0.011*"time" + 0.011*"nigh
t" + 0.011*"heavi" + 0.010*"love" + 0.008*"tweet"
Topic: 4 Word: 0.030*"short" + 0.022*"bodi" + 0.021*"good" + 0.015*"hair" + 0.014*"guy" + 0.013*"your" + 0.012*"thin
g" + 0.011*"life" + 0.011*"wanna" + 0.011*"mind"
Topic: 5 Word: 0.027*"short" + 0.022*"tall" + 0.018*"hair" + 0.016*"think" + 0.015*"bodi" + 0.012*"like" + 0.011*"mor
n" + 0.011*"heavi" + 0.010*"skinni" + 0.010*"dark"
Topic: 6 Word: 0.021*"short" + 0.018*"love" + 0.014*"skin" + 0.013*"bodi" + 0.011*"ugli" + 0.011*"dark" + 0.011*"heav
i" + 0.010*"friend" + 0.010*"have" + 0.010*"go"
Topic: 7 Word: 0.038*"bodi" + 0.033*"ugli" + 0.024*"short" + 0.014*"know" + 0.014*"like" + 0.013*"dark" + 0.012*"goo
d" + 0.012*"love" + 0.011*"time" + 0.010*"dont"
```

Looking at the words, I classified the topics as following:

Topic 0: not body
Topic 1: body
Topic 2: body
Topic 3: body
Topic 4: body
Topic 5: body
Topic 6: not body
Topic 7: not body

2. As a document can belong to multiple topics, next I looked at what topic is the dominant topic in a document.

	Document_No	Dominant_Topic	Topic_Perc_Contrib	Keywords	Text
0	0	1	0.6805	skin, heavi, feel, dark, skinni, short, rain, ...	[bodi, feel, itchi, like]
1	1	6	0.5611	short, love, skin, bodi, ugli, dark, heavi, fr...	[fall, asleep, hear, traci, girl, bodi, heart,...
2	2	1	0.6978	skin, heavi, feel, dark, skinni, short, rain, ...	[hate, peopl, diss, band, trace, clearli, ugli]
3	3	2	0.5622	short, like, love, bodi, dark, watch, week, th...	[one]
4	4	7	0.7421	bodi, ugli, short, know, like, dark, good, lov...	[thecoolestout, ehhh, dont, weather, gonna, tu...
5	5	4	0.6638	short, bodi, good, hair, guy, your, thing, lif...	[mind, bodi, sever, protest, quotget, upquot, ...
6	6	2	0.6143	short, like, love, bodi, dark, watch, week, th...	[geez, busi, afternoon, meet, email, meet, ema...
7	7	0	0.7023	dark, look, short, bodi, skin, go, miss, work,...	[time, posterior, lose, articul, creak, run, d...
8	8	1	0.6755	skin, heavi, feel, dark, skinni, short, rain, ...	[wait, skinni, vega, hungri]
9	9	5	0.4389	short, tall, hair, think, bodi, like, morn, he...	[wait, airport, ride, harass, tri, sell, ugli,...

For example, the dominant topic of document 0 is 1, that of document 1 is 6, that of document 4 is 7 and so on. I used this analysis to categorize whether the document was 'body' or 'not body' based on my classification of topics above.

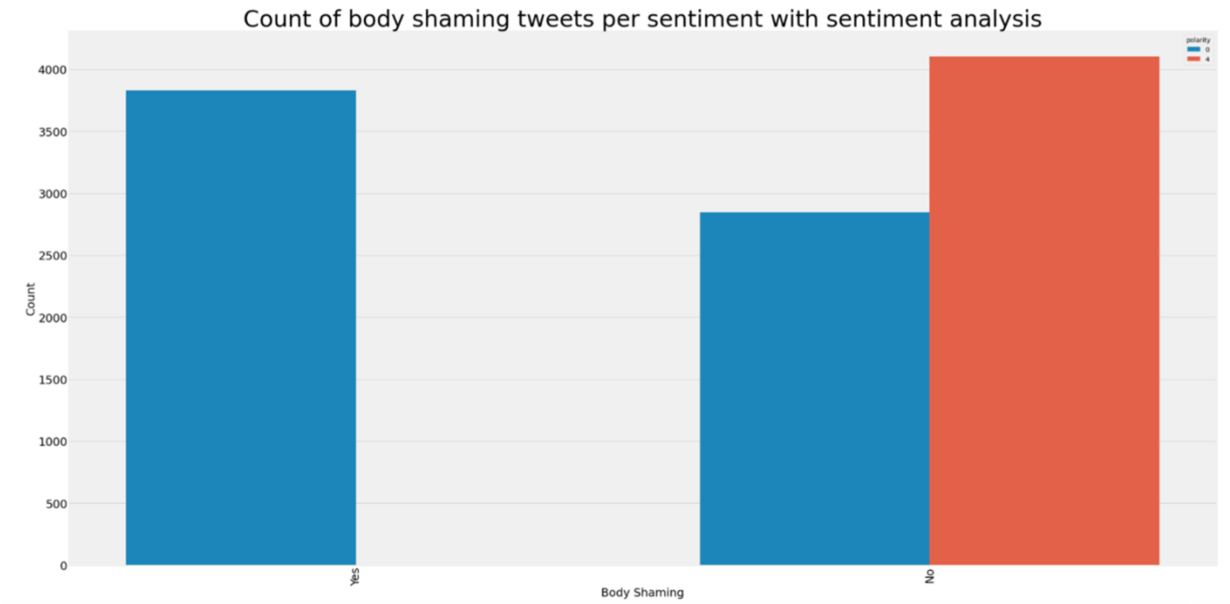
	polarity	text	text_2	Document_No	Dominant_Topic	Topic_Perc_Contrib	Keywords	Text	Topic
0	0	my whole body feels itchy and like its on fire	my whole body feels itchy and like its on fire	0	1	0.6805	skin, heavi, feel, dark, skinni, short, rain, ...	[bodi, feel, itchi, like]	body
1	0	Falling asleep. Just heard about that Tracy gi...	Falling asleep Just heard about that Tracy gir...	1	6	0.5611	short, love, skin, bodi, ugli, dark, heavi, fr...	[fall, asleep, hear, traci, girl, bodi, heart,...	not body
2	0	i really hate how people diss my bands! Trace...	i really hate how people diss my bands Trace ...	2	1	0.6978	skin, heavi, feel, dark, skinni, short, rain, ...	[hate, peopl, diss, band, trace, clearli, ugli]	body
3	0	why is it always the fat ones?!	why is it always the fat ones	3	2	0.5622	short, like, love, bodi, dark, watch, week, th...	[one]	body
4	0	@thecoolestout Ehhh don't. Weather's gonna tak...	thecoolestout Ehhh dont Weathers gonna take a ...	4	7	0.7421	bodi, ugli, short, know, like, dark, good, lov...	[thecoolestout, ehhh, dont, weather, gonna, tu...	not body

For example, in the table above, documents 0, 2, 3 are classified as 'body' and documents 1,4 are classified as 'not body'.

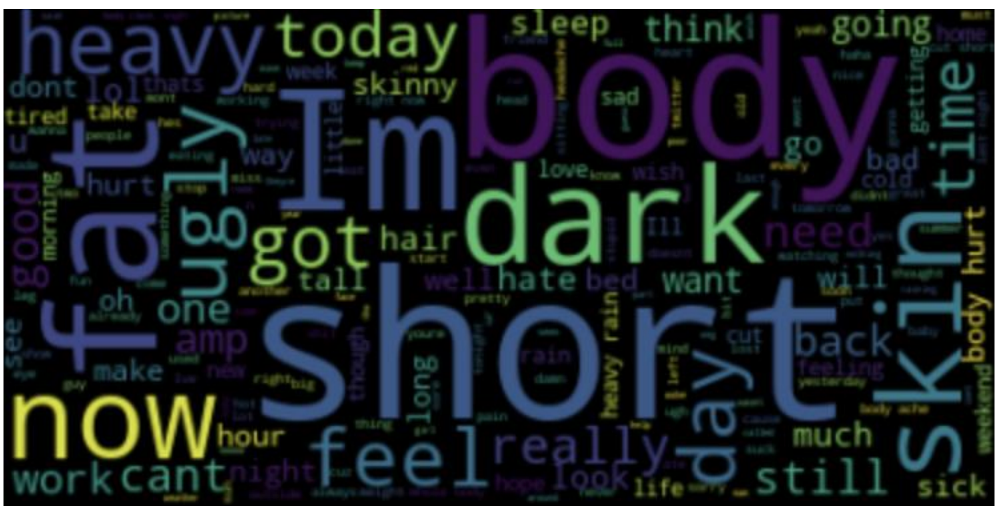
3. Finally, if the previously analyzed sentiment of the tweet is 0 (negative) and the topic is 'body', it is classified as 'Body Shaming' and the if the sentiment is 2 (neutral) and 4 (positive) and the topic is 'body', it is classified as 'Not Body Shaming'.

	polarity	text	text_2	Document_No	Dominant_Topic	Topic_Perc_Contrib	Keywords	Text	Topic	Body_Shaming
0	0	my whole body feels itchy and like its on fire	my whole body feels itchy and like its on fire	0	1	0.6805	skin, heavi, feel, dark, skinni, short, rain, ...	[bodi, feel, itchi, like]	body	Yes
1	0	Falling asleep. Just heard about that Tracy gi...	Falling asleep Just heard about that Tracy gir...	1	6	0.5611	short, love, skin, bodi, ugli, dark, heavi, fr...	[fall, asleep, hear, traci, girl, bodi, heart,...	not body	No
2	0	i really hate how people diss my bands! Trace...	i really hate how people diss my bands Trace ...	2	1	0.6978	skin, heavi, feel, dark, skinni, short, rain, ...	[hate, peopl, diss, band, trace, clearli, ugli]	body	Yes
3	0	why is it always the fat ones?!	why is it always the fat ones	3	2	0.5622	short, like, love, bodi, dark, watch, week, th...	[one]	body	Yes
4	0	@thecoolestout Ehhh don't. Weather's gonna tak...	thecoolestout Ehhh dont Weathers gonna take a ...	4	7	0.7421	bodi, ugli, short, know, like, dark, good, lov...	[thecoolestout, ehhh, dont, weather, gonna, tu...	not body	No

Out of the 10782 tweets, 3830 are classified as 'Body Shaming' and 6952 are classified as 'Bot Body Shaming'.



I also generated a word cloud of tweets that are classified as 'Body Shaming'.



Discussion

In this project, I used the techniques of Natural Language Processing to identify body shaming tweets. I combined NLP with the previously done sentiment analysis to classify tweets as body shaming or not shaming.

Example of tweets that have been identified as body shaming that fits the purpose of the project. These were fewer in number than I expected.

- "I ate lunch outside for no more then minutes today And Im pink Stupid burnable skin"
- "why is it always the **fat ones**"

- “**Feeling like a fat chick** barely fitting in her clothes”
- “at my gfs **feeling pretty fat** have to go home and get work clothes”
- “Chuck Liddell is so out of shape **its** not even funny the biggest belly fat Ive seen in the Octagon”
- “i hate my pale scottish skin burnt already lol”
- “aplusk I agree But I gain those few pounds right back My skinny euphoria is always shortlived”
- “im so fat wanna loose wait i wanna be lb”
- “my body isnt delicious”
- “Got a red bathing suit I am still not skinny enough for a bikini httpyfrogcometmj”
- “i hate how dark my hair is”
- “great not the only one here anymore ugly fat lady is here w ugly gay friendsoh wait they left”
- “what an ass whole i am super pissed at his dumb ugly skinny self”
- “Flatironing hair I hate my ugly curls”
- “A show about fat people dancing Because everything is funnier when fat people do it”
- “fighting with fat girls They are so mean Fuck it Eat mcdonalds”

Example of tweets that have been identified as body shaming that is not really body shaming. A lot of the tweets in this group were about people feeling sick.

- “Gosh Its taking my body a long time to bounce back from my Atlanta trip Just way too much stress Been flaring ever since”
- “new bad habit picking at the skin on my upper arms I am bleeding”
- “saw a skinny kitten with its mother I gave some food soooorrryyy kitty TT”
- “its really dark outside and its only PM its going to rain”
- “I think imma lay in bed today my body is hurting”
- “my whole body feels itchy and like its on fire”

Example of tweets that have not been identified as body shaming that could potentially have been identified as such:

- “Jnavolio ah thanksive been workin out hard for like months and you go and call me fat”
- “kellyerin they are fat and old and nasty and flirt with meGAHHHH”
- “ashleyblah but i am fat you look good with bangs btw”
- “im a little disappointed with amber and the other skinny lil white girl at dunkin this moening”
- “Im so ugly i know”
- “Other models are making me feel short Imagine”
- “feels like my hair is too dark to wear anything nice to work”
- “im sooooooo fat what i doing”
- “i hate my body so many scars and bumps”
-

Example of tweets that were correctly not identified as body shaming:

- “viaHourt ahhh i see i see im in bed watching ugly betty lol”
- “My friends short story is going to press Congrats Melody”

- "im out walking and its cold and dark and i heard something behind me S ahhhhhhh"
- "BHGRESherry Looks like U are good to go That was ugly stuff"

Limitations

Identifying body shaming tweets is by no means a perfect endeavor. I acknowledge that there are some limitations to the process.

1. First, reproducibility could be an issue here. Some steps of the classification and filtering is based on my knowledge, research, or analysis. For example, I used my informed judgement while filtering the dataset for tweets containing words/phrases/sentences related to body based on my interaction with social media platforms. It was hard to find an already existing dictionary or list of words that served the purpose.
2. Similarly, the topics that were generated after topic modeling were classified as 'body' or 'not body' based on my judgement of what the dominant words in the topic were. A different person would probably classify those topics differently.
3. There wasn't a lot of scope for hyperparameter tuning to make the model more effective. I used coherence score which is one of the few suitable available for LDA topic modeling. The discussions prove helpful in this regard. The discussion above somewhat serves as a correlation matrix.
4. Another major limitation is that this model doesn't have the potential to identify sarcasm.