# Analyzing Twitter's Trending Topics

Virgilia Antonucci

Music · Trending

**BETTER THAN REVENGE**

30.5K posts

Trending in United States

**Courtney**

Trending with Justine, Justine

Entertainment · Trending

**Fassbender**

2,224 posts

Politics · Trending

**Jill Stein**

3,591 posts

Trending

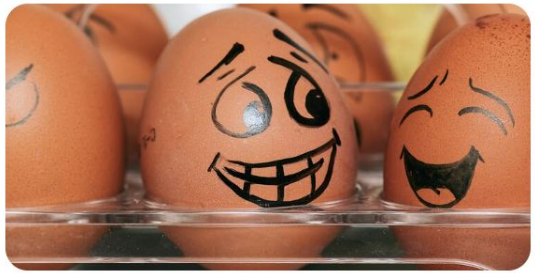
**The Social Network**

4,656 posts

# Data

## Emotion Classification NLP

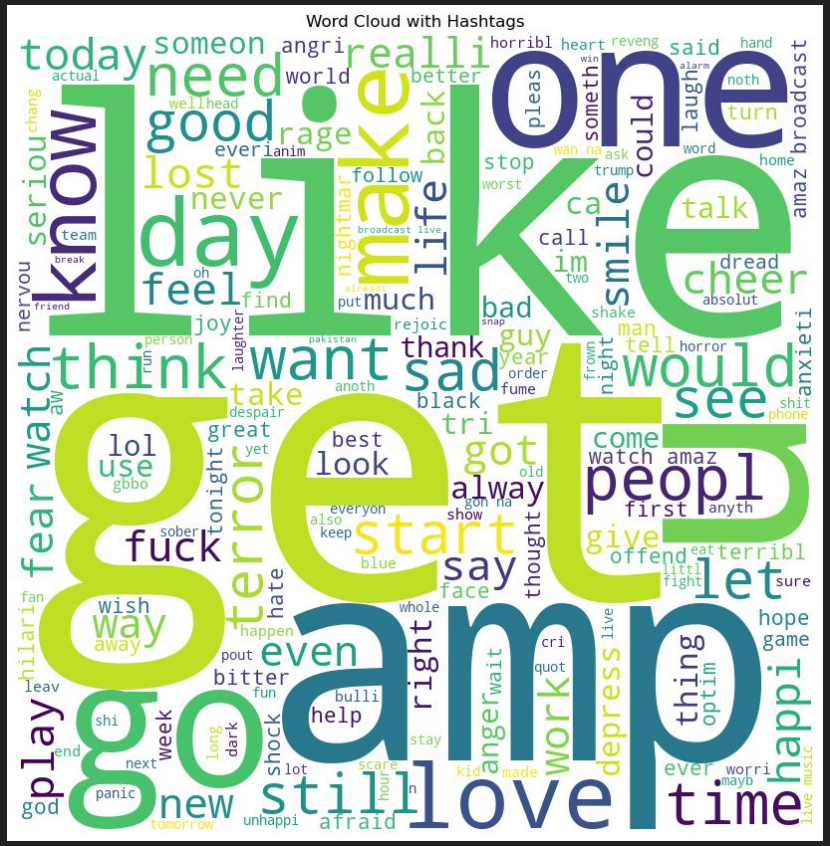
Tweets classified based on 4 emotions - joy, sadness, anger and fear.

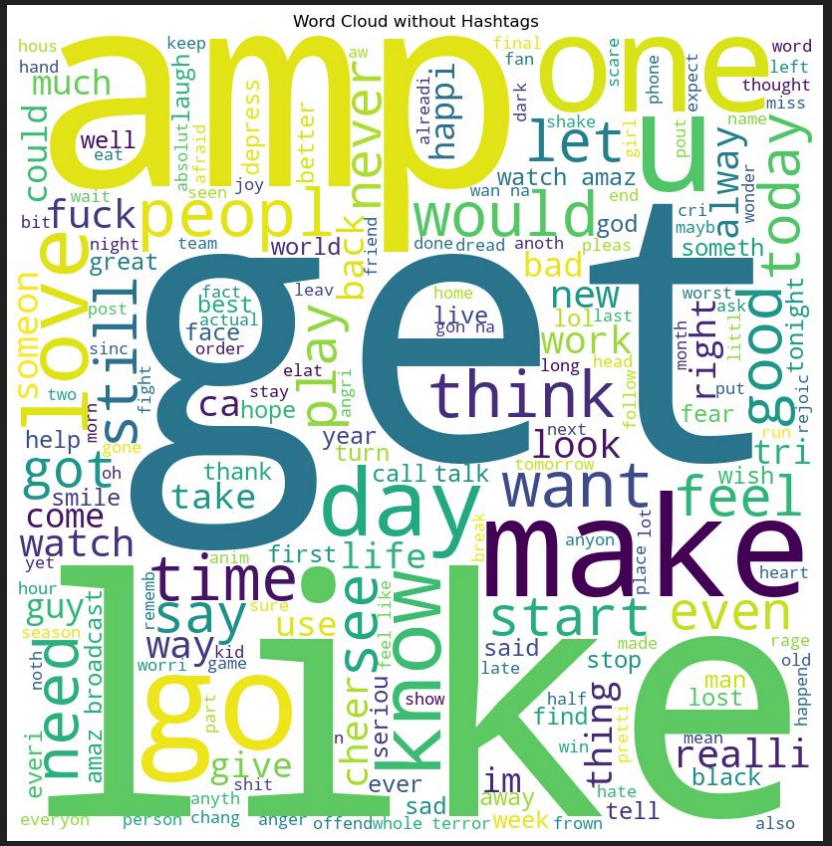


# Steps

- Lowercasing
- Removing Mentions
- Tokenizing
- Stemming
- Vectorizing

We end up with two sets of data. One is tweets with hashtags and one is tweets without.


Word Cloud with Hashtags
one
like
get
go
amp
love
time
know
day
make
want
think
sad
see
peopl
start
fuck
lol
use
way
even
still
lost
good
need
today
someon
realli
feel
look
come
got
let
watch
terror
fear
play
happi
thing
work
depress
anger
life
cheer
smile
talk
dread
rage
hope
game
wish
bitter
shock
help
panic
afraid
unhappi
tomorrow
broadcast
amaz
anxieti


Word Cloud without Hashtags
amp
one
get
like
make
love
go
time
think
day
want
let
know
peopl
would
still
need
start
feel
look
come
watch
take
play
life
see
say
way
even
realli
today
good
work
cheer
thing
fuck
never
always
happi
broadcast
amaz
tri
right
hope
smile
give
use
sad
black
tell
week
frown
anger
offend
whole
terror
person
also

# Most Common Words

Most common words with hashtags:

like 253

get 227

amp 182

go 164

make 155

one 149

day 144

watch 141

love 138

time 130

Most common words without hashtags:

like 250

get 227

amp 182

go 164

make 154

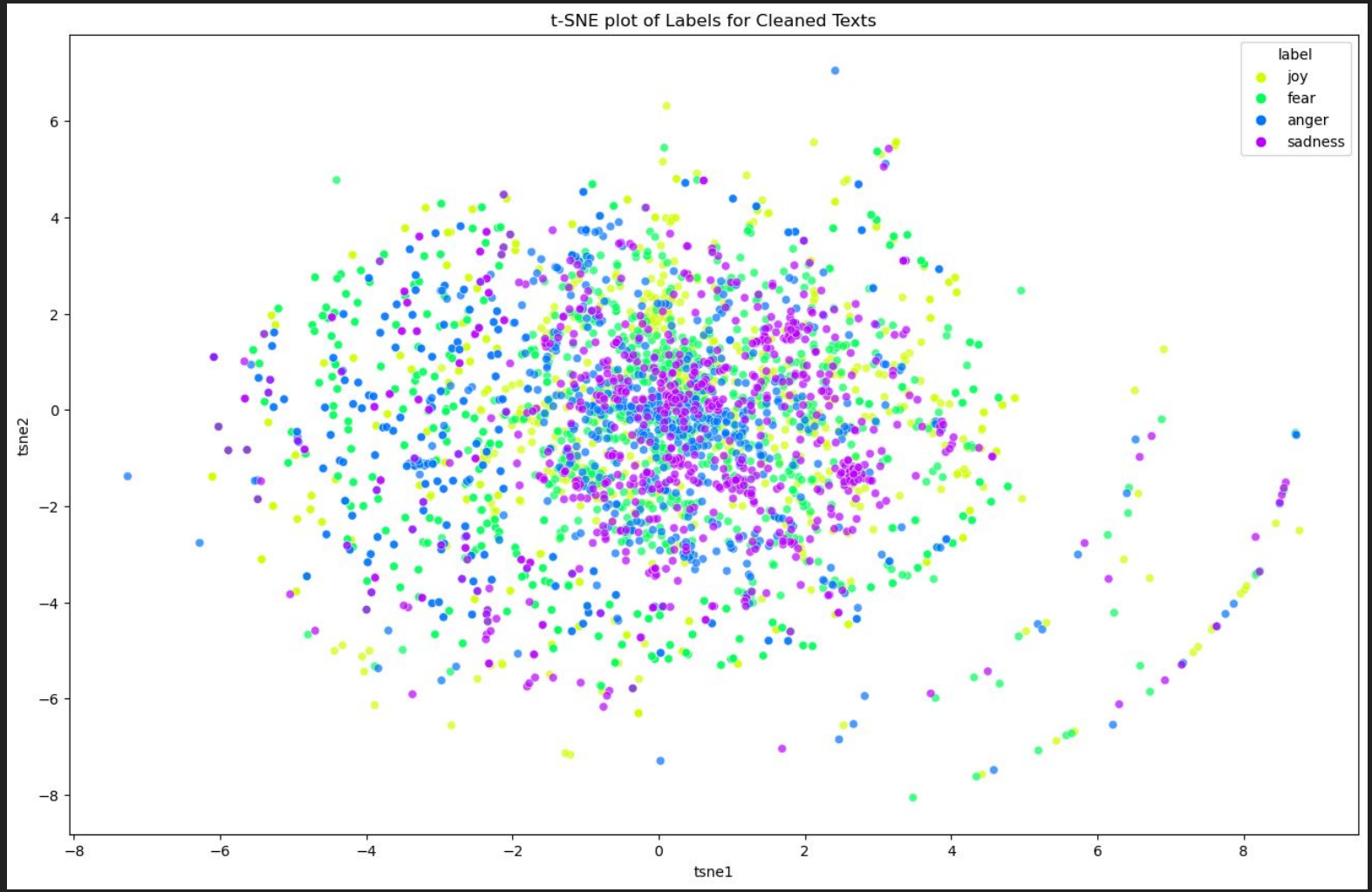
one 149

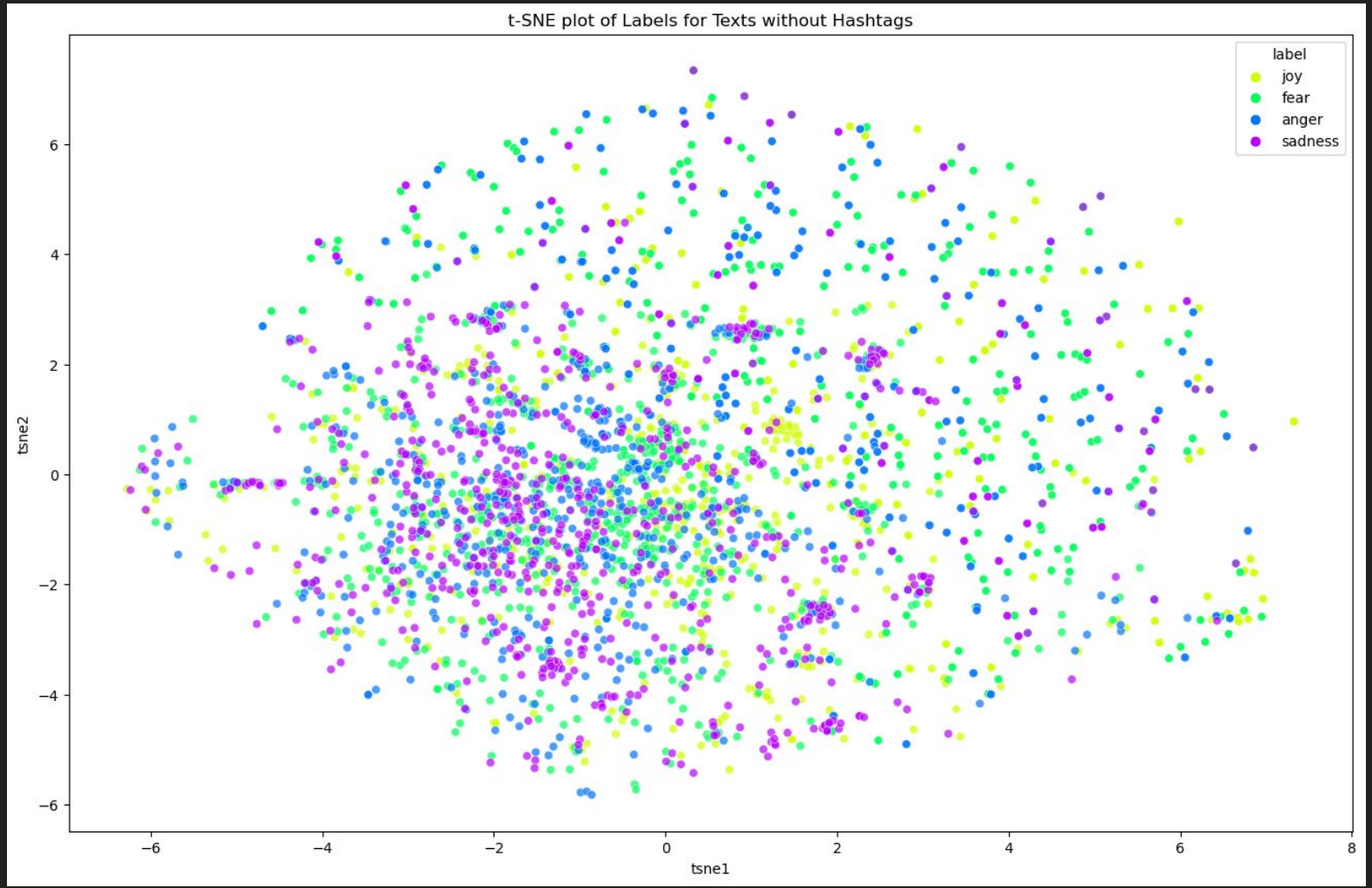
day 144

watch 140

time 127

know 126


t-SNE plot of Labels for Cleaned Texts
label
joy
fear
anger
sadness
tsne1
tsne2
6
4
2
0
−2
−4
−6
−8
−8
−6
−4
−2
0
2
4
6
8


t-SNE plot of Labels for Texts without Hashtags
label
joy
fear
anger
sadness
tsne1
tsne2
6
4
2
0
−2
−4
−6
−6
−4
−2
0
2
4
6
8

# Support Vector Machine with GridSearch

```
Classification Report:
               precision    recall  f1-score   support

       anger       0.86      0.85      0.86       154
        fear       0.85      0.91      0.88       223
         joy       0.92      0.92      0.92       185
     sadness       0.82      0.76      0.79       161

    accuracy                           0.87       723
   macro avg       0.86      0.86      0.86       723
weighted avg       0.87      0.87      0.87       723

--------------------------------------------------

Fitting 5 folds for each of 36 candidates, totalling 180 fits
Best parameters for Texts without Hashtags:
{'svm__C': 10, 'svm__kernel': 'linear', 'tfidf__max_df': 0.5, 'tfidf__ngram_range': (1, 2)}
Results for Texts without Hashtags:
Accuracy: 0.7427385892116183

Classification Report:
               precision    recall  f1-score   support

       anger       0.73      0.74      0.74       154
        fear       0.71      0.78      0.74       223
         joy       0.86      0.85      0.86       185
     sadness       0.65      0.57      0.61       161

    accuracy                           0.74       723
   macro avg       0.74      0.73      0.74       723
weighted avg       0.74      0.74      0.74       723

--------------------------------------------------
```

# Conclusion:

# Hashtags are a must!

# Next Step

- Expand Sentiment analysis to more emotions to handle more nuance

# Thank you.

LinkedIn GitHub