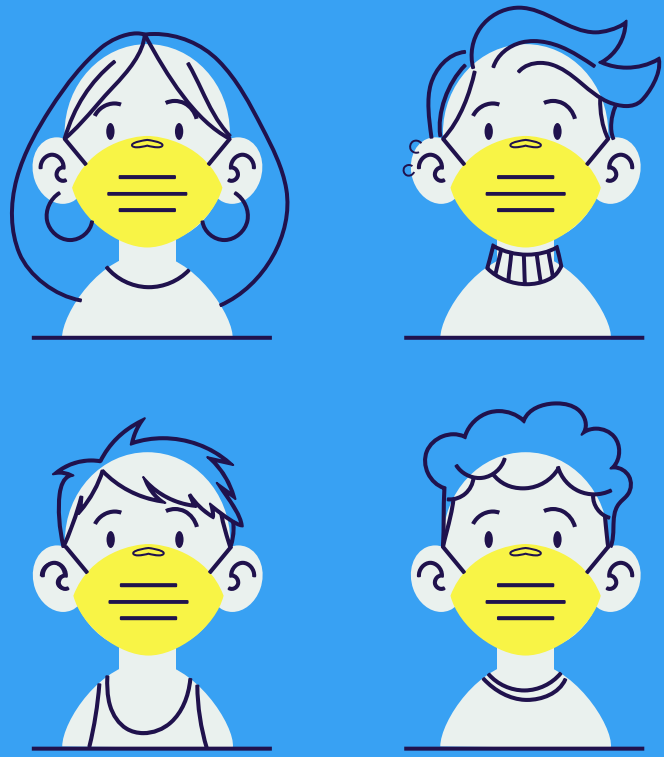
# #MASKS THROUGHOUT COVID-19

Twitter Sentiment Analysis



# EXPERT SENTIMENT OVER TIME

"Seriously people- STOP BUYING MASKS! They are NOT effective in preventing general public from catching #Coronavirus..."

@Surgeon_General
2020.02.29

"...wearing a mask is completely safe. As an anesthesiologist, I wear a mask all day long to protect my patients..."

@Surgeon_General
2020.05.23

-0.7536 **Vader Sentiment Score** 0.7257

## DATE RANGE

January 1, 2020
-
May 31, 2020

## SEARCH TERMS

(mask OR masks)
+
(covid OR covid19 OR corona OR coronavirus)

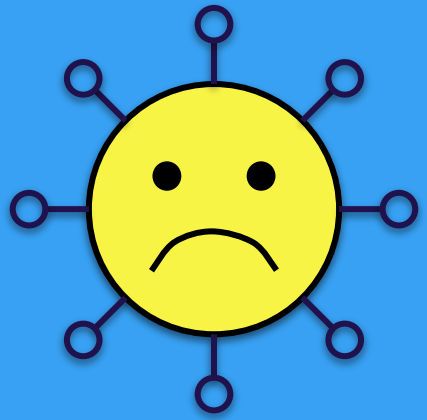
## TOTAL TWEETS

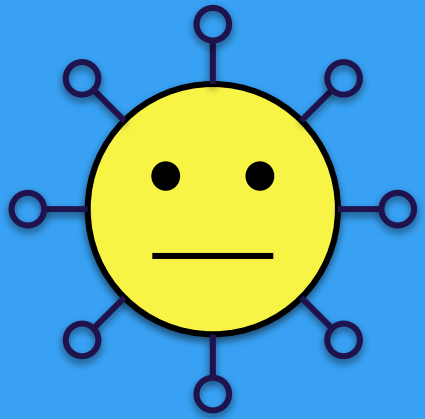
5,000 tweets/day
(580,000)

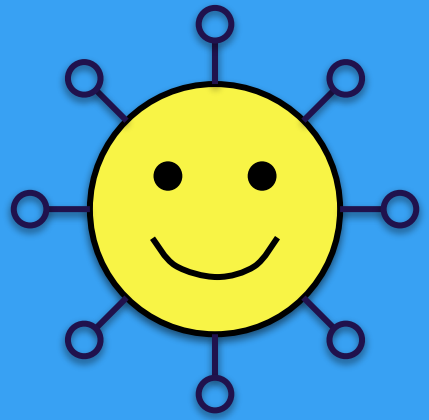
## OTHER INFO

Num of likes
Num of replies
Num of retweets

# TARGET!

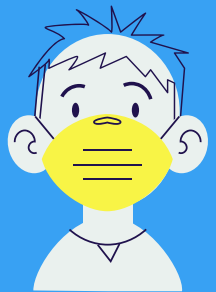


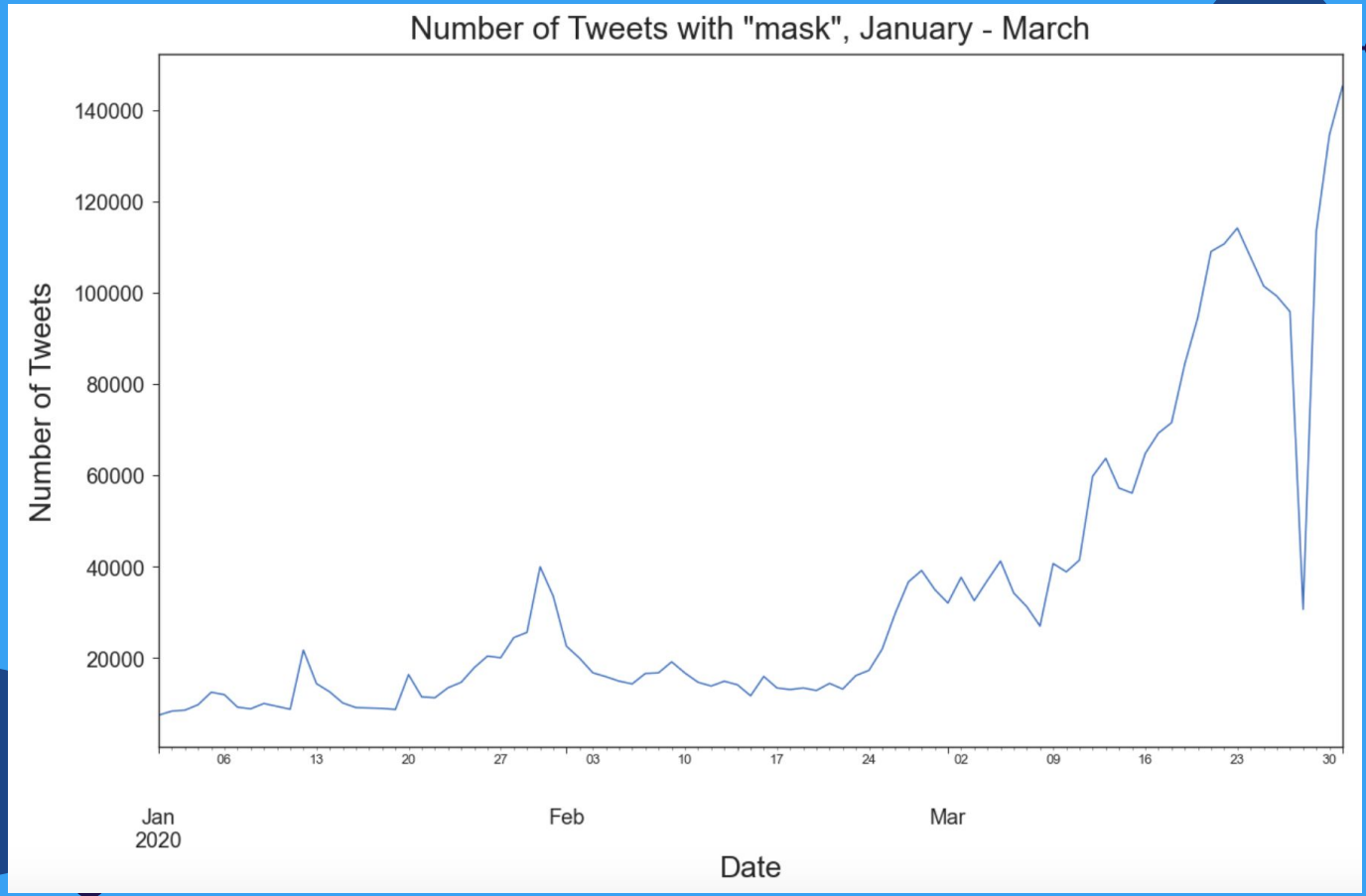


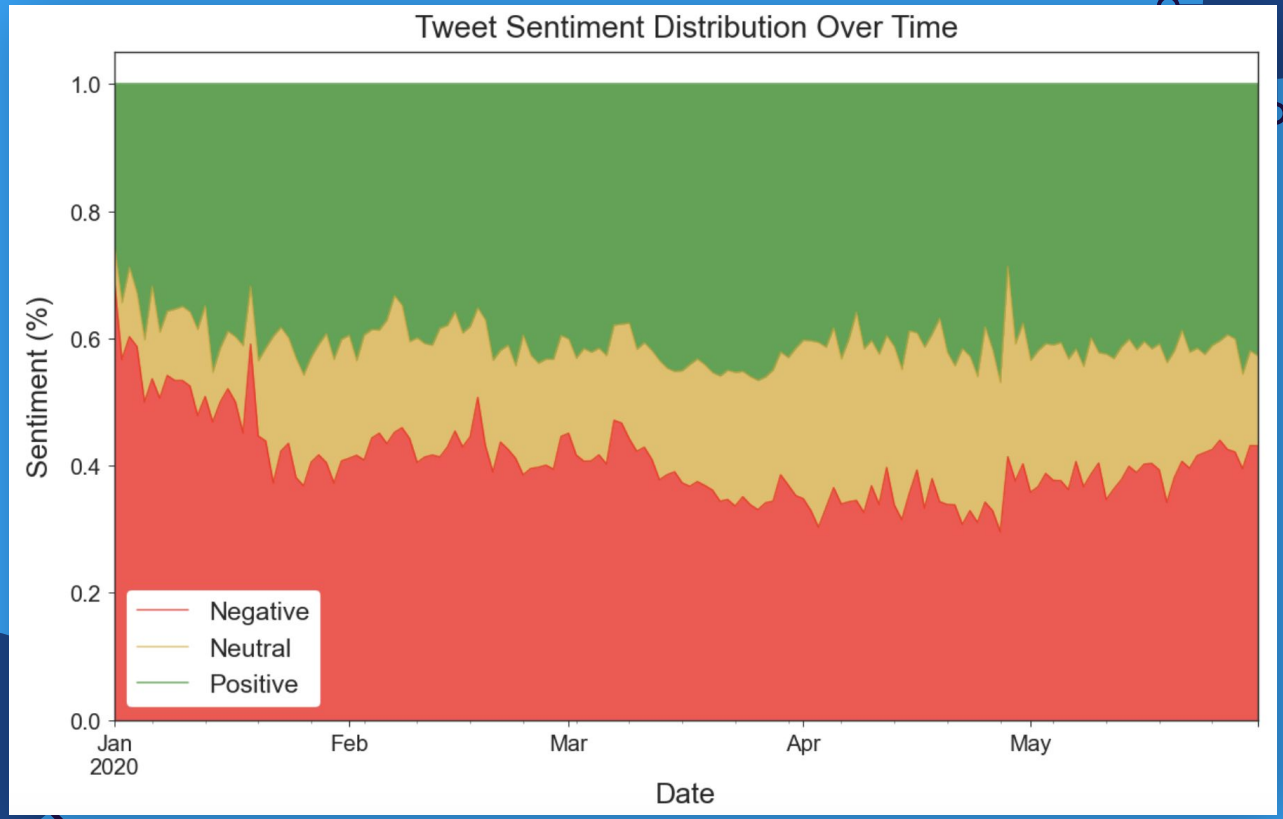


# ROADMAP

1. Scrape and Prep Data
2. Target Variable Creation
3. Exploratory Data Analysis
4. Feature Engineering
5. Prediction Modeling

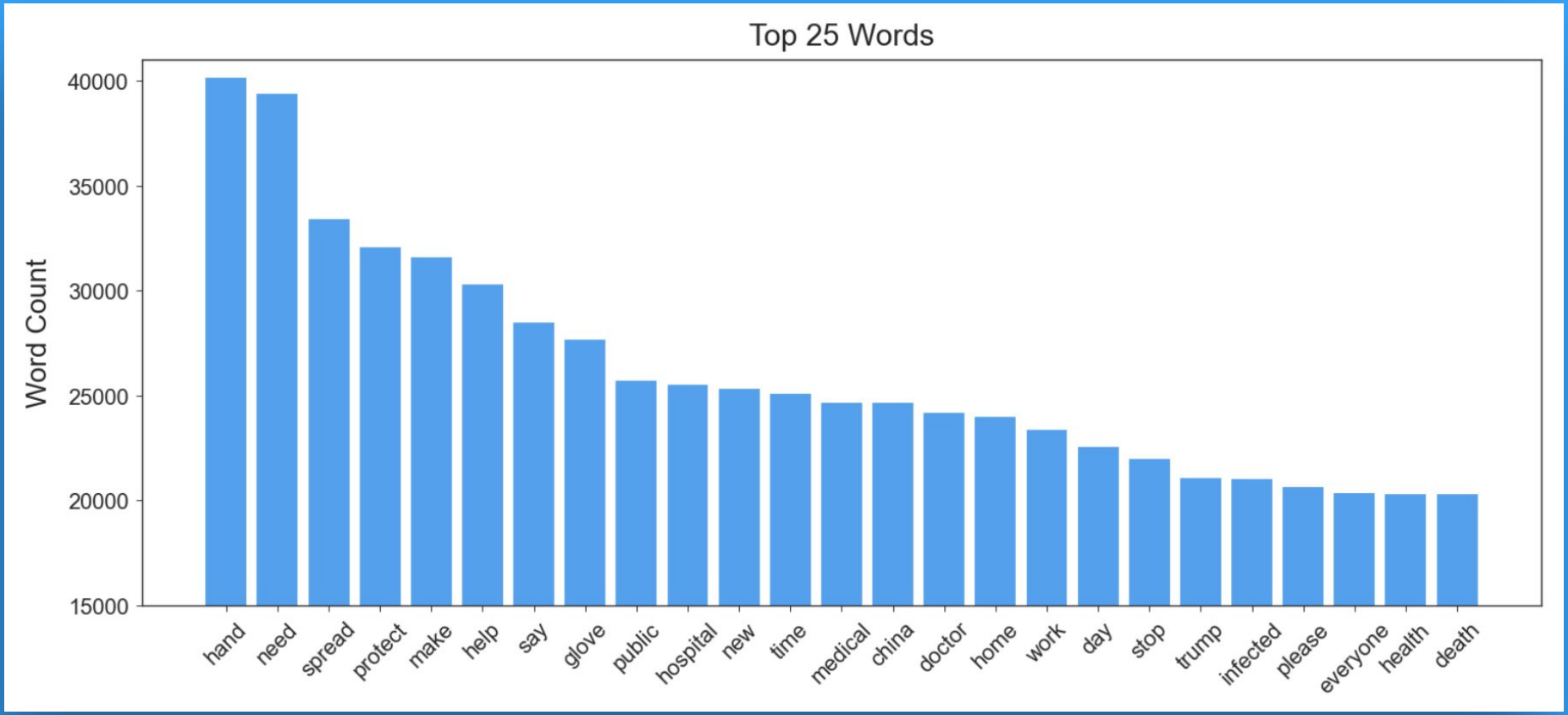


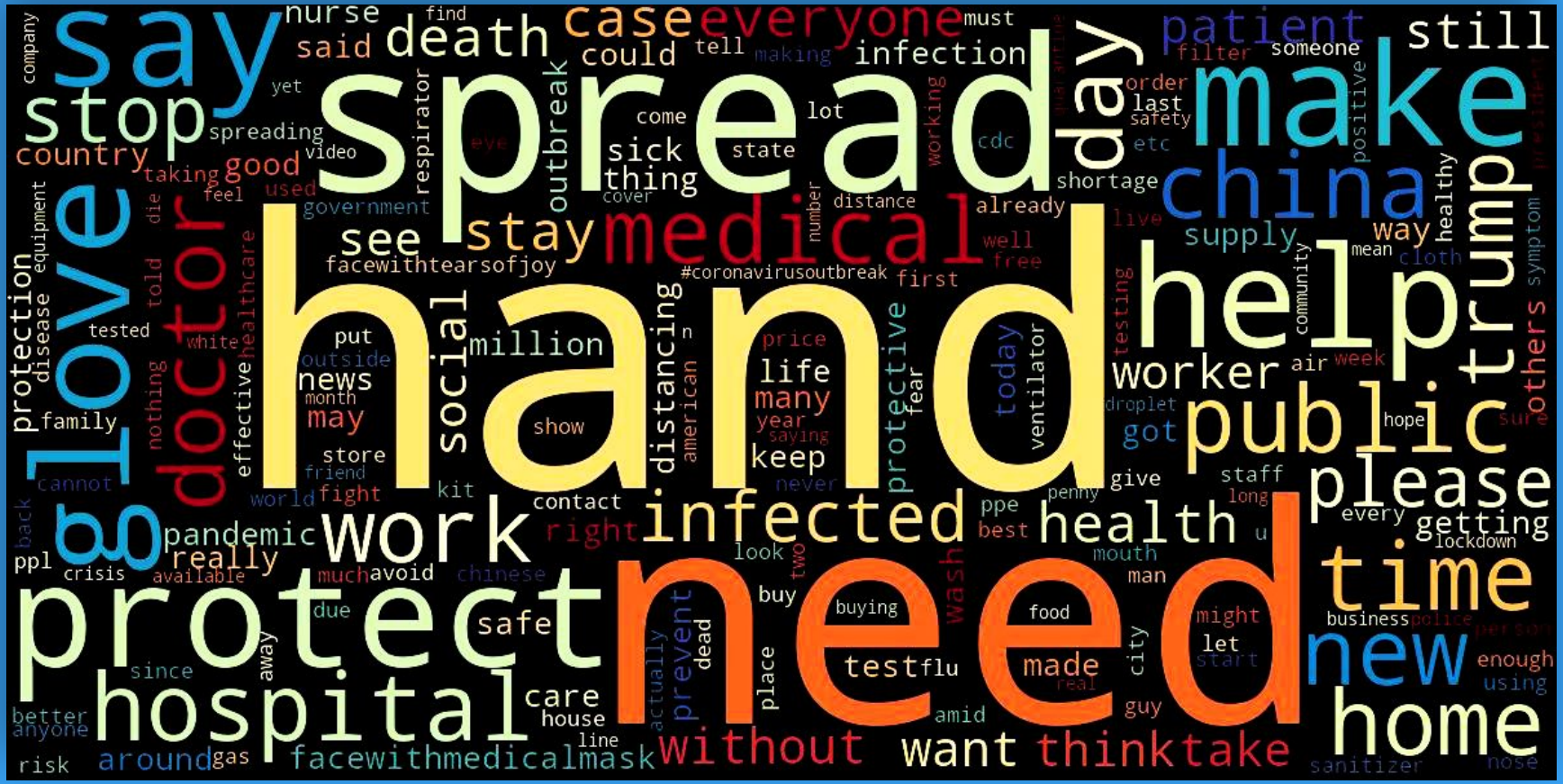

Number of Tweets with "mask", January - March
Number of Tweets
140000
120000
100000
80000
60000
40000
20000
06
13
20
27
03
10
17
24
02
09
16
23
30
Jan
2020
Feb
Mar
Date


Tweet Sentiment Distribution Over Time
Sentiment (%)
1.0
0.8
0.6
0.4
0.2
0.0
Negative
Neutral
Positive
Jan
2020
Feb
Mar
Apr
May
Date

# NLP FEATURES AND ANALYSIS

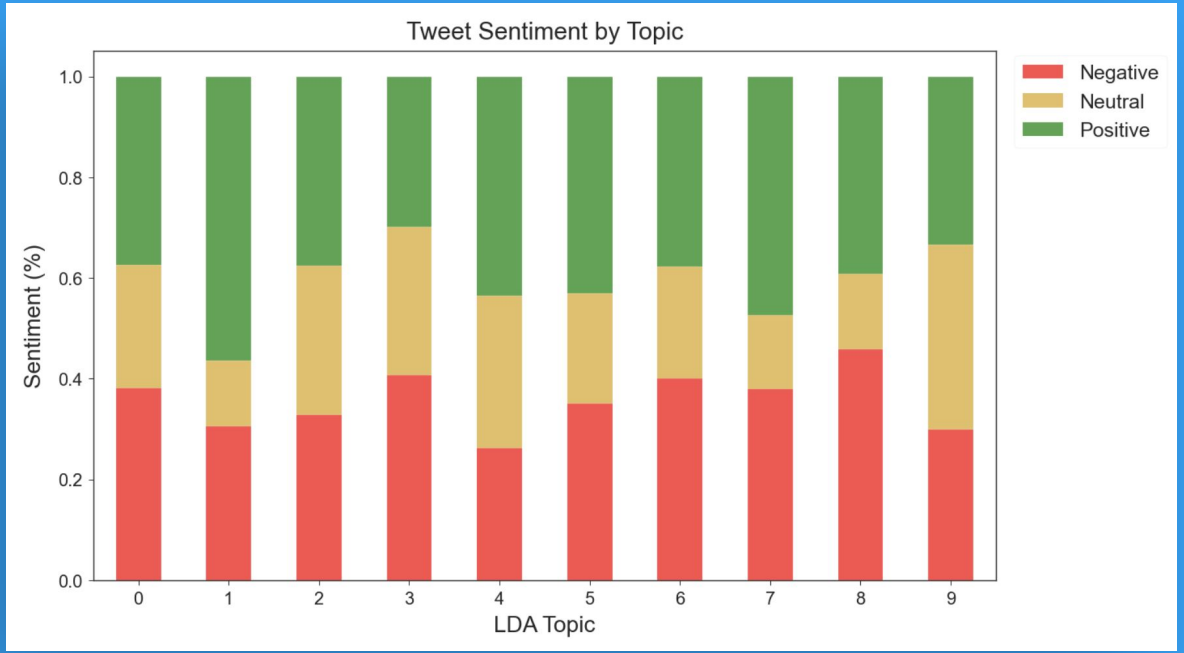
Text cleaning, topic modeling, subjectivity scoring


Top 25 Words
Word Count
40000
35000
30000
25000
20000
15000
hand
need
spread
protect
make
help
say
glove
public
hospital
new
time
medical
china
doctor
home
work
day
stop
trump
infected
please
everyone
health
death



hand
need
spread
protect
say
make
help
public
time
new
home
hospital
work
infected
health
please
trump
china
medical
stay
doctor
love
stop
day
patient
still
everyone
case
death
nurse
said
could
infection
country
good
sick
thing
see
facewithtearsofjoy
#coronavirusoutbreak
million
social
distancing
protective
worker
supply
way
cloth
others
healthy
positive
order
last
safety
etc
shortage
already
well
free
first
price
life
many
year
keep
today
ventilator
fear
american
show
put
outside
news
month
may
store
friend
fight
world
kit
contact
right
pandemic
really
much
avoid
chinese
safe
care
house
line
facewithmedicalmask
around
gas
risk
better
anyone
since
away
due
actually
prevent
dead
place
buy
buying
test
flu
made
food
city
amid
guy
might
let
start
business
police
enough
using
sanitizer
nose
mouth
man
look
two
wash
best
ppe
penny
give
staff
long
every
getting
lockdown
hope
sure
air
week
droplet
got
testing
community
mean
live
number
distance
cover
state
lot
come
eye
respirator
outbreak
tell
making
must
filter
someone
positive
president
company
yet
spreading
video
taking
feel
used
government
die
equipment
protection
disease
tested
white
healthcare
told
nothing
effective
family
cannot
back
ppl
crisis
available
back

0. Healthcare workers, hospitals
1. Social distancing
2. Protesting, lockdowns
3. Government, health organizations
4. Spreading the virus
5. Emojis, swear words
6. COVID19 statistics
7. Preventing infection
8. General opinions
9. Riots, BLM




# TOPIC MODELING

# PREDICTION MODEL FEATURE IMPORTANCE

| | |
|---|---|
| **IMPORTANT FEATURES** | '#fight' 'incompetence' |
| **ACCURACY** | 70.1% |
| **FI SCORE** | 69.9% |

Naive Bayes

| | |
|---|---|
| **IMPORTANT FEATURES** | Subjectivity score<br>Number of likes<br>'protect'<br>'help' |
| **ACCURACY** | 72.3% |
| **FI SCORE** | 72.4% |

Decision Tree

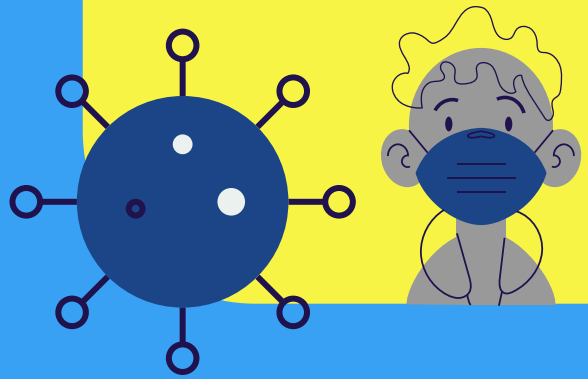
# TAKEAWAYS AND FUTURE CONSIDERATIONS

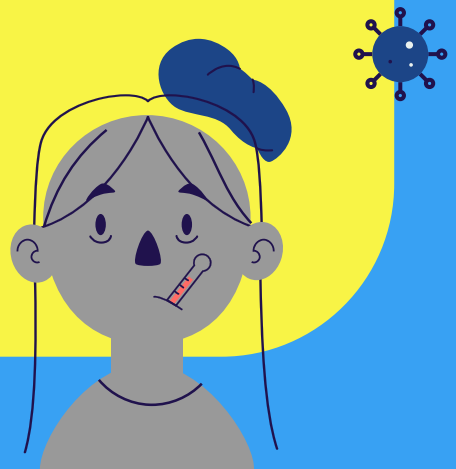
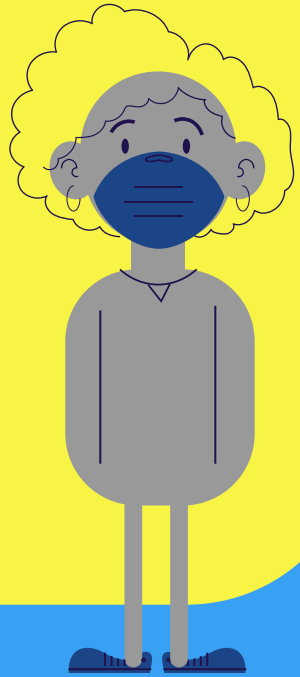
Overall sentiment is fairly divided

Unexpected feature importance

Try a deep learning model

Opinion mining toward “mask(s)”





# QUESTIONS?