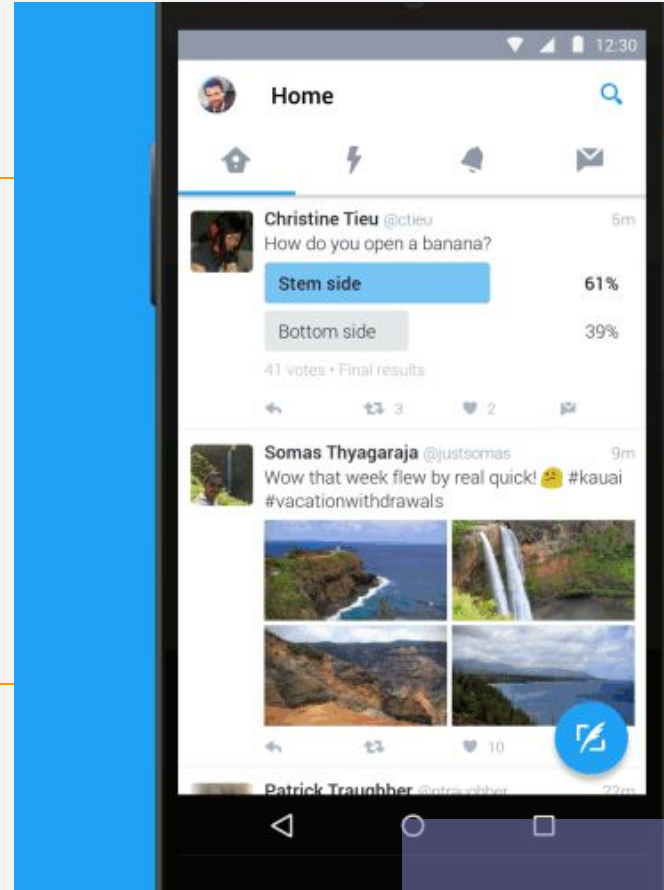


SOCIAL MEDIA DISASTER ALERT SYSTEM

Utilizing social media to alert about new
disasters and their nature





PROBLEM STATEMENT

FEMA has requested a demo of an emergency alert system for a large metropolitan area based on social media "reporting." Twitter provides a live feed to what's happening around the Greater New York region before many emergencies even get reported. We need to utilize this of-the-minute information to alert FEMA of a potential emergency as well as where it's approximately located.

TWEETS

KEY TERMS

#Blackout

Key terms like "blackout" are assessed and sited as a potential emergency



GLYNEWS @NewYork_SPIN · Jul 14

A massive power outage in New York City left streets around Times Square and Broadway — some of the busiest parts of the city — without electricity.

#nycpoweroutage #Blackout

#NYC #NYCblackout #blackoutnyc #Blackout2019 #NewYork #manhattan



0:11 37.8K views



23



234



654



LIVE COVERAGE

Assessing Emergencies

Twitter provides live feedback about what's happening around NYC, often with images or video



PROCESS

DATA GATHERING

We gathered several large "emergency" datasets for training and gathered "live" tweets from the Twitter API developer app

EDA

We reviewed and cleaned the data through and vectorizing using lemmatizing

MODELING

We performed multiple grid searches and tuned the model's parameters to continually improve our scores

01

04

02

05

03

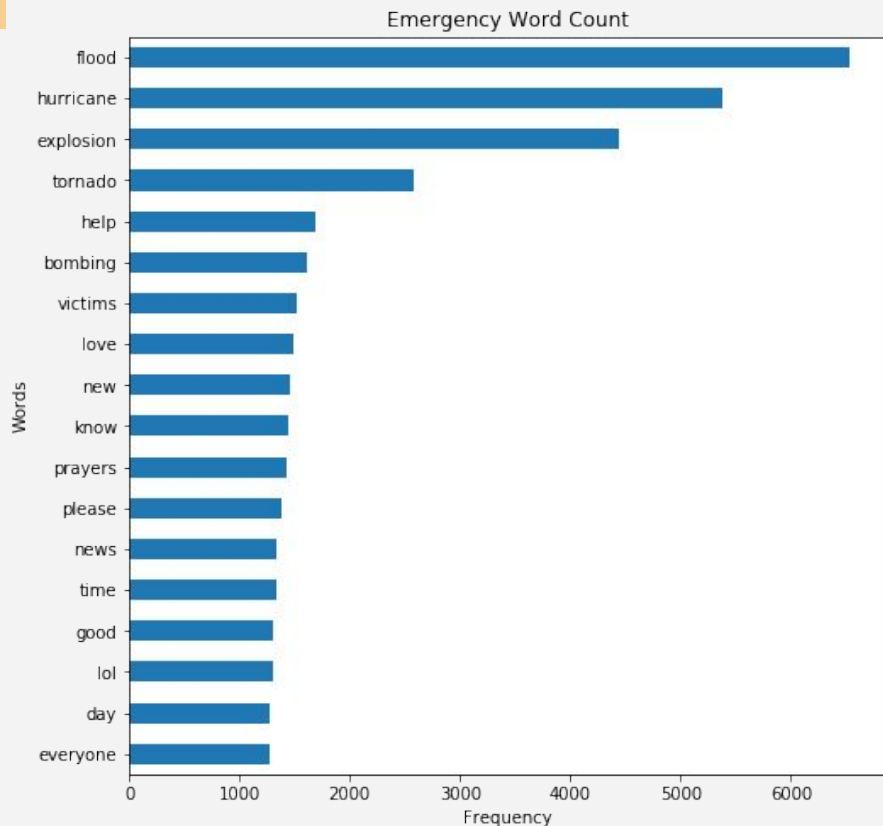
ANALYSIS

Once we felt confident in our train/test scores we ran the "live" tweets through the same grid search parameters to label them as emergencies or not

ALERT SYSTEM

When tweets trigger the emergency terms as learned by our models, alerts will show up on a map of the Greater New York City region with approximate locations of the tweets

DATA GATHERING / EDA



DATA

Data was pulled from a Kaggle set to train models while we created a Twitter developer account to eventually pull the live tweets

KEY TERMS

We trained the model with a lexicon of 380 emergency terms including terms like "flood", "blackout" and "help"

FINDINGS

The chart to the left shows the top unique terms in our training set

SCORES / ACCURACY

94.88

Test Score

LOGISTIC REGRESSION

CountVectorizer performed better, 1500 features and an ngram range of (1, 2) and stop words

93.65

Test Score

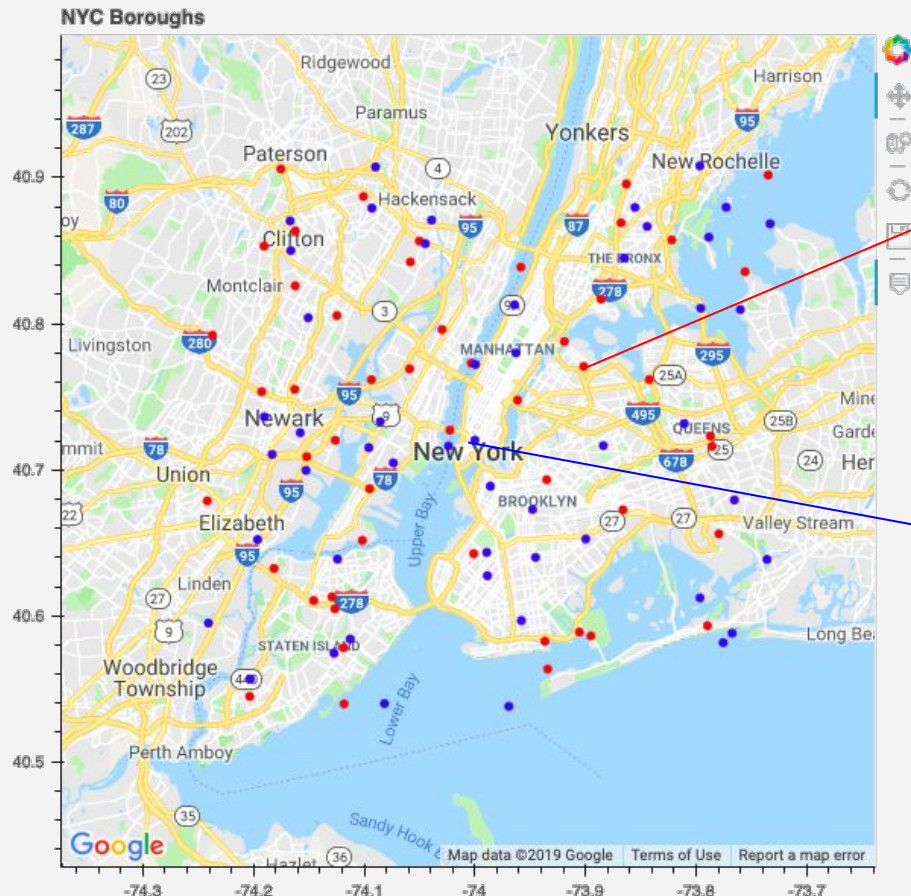
RANDOM FOREST

CountVectorizer performed better with an ngram range (1, 1), including stop words

CONFUSION MATRIX

True Positives	6,530	375	False Negatives
False Positives	394	7,722	True Negatives

ANALYSIS & ALERTS



EMERGENCY



Samarjeet
@sjsgrewal

@AmitShah @narendramodi @RahulGandhi
@DeepikaBhardwaj @priyankagandhi @WIONews
@CNNnews18 SOMEBODY PLEASE SEND HELP! I AM
NOT SAFE! #SOS #HELP #Emergency Did EVERYTHING
asked from me. Just wants to kill me tonight. Won't stop
with the deliberate physical harm! Even deboarded a flight

1:53 PM · Jul 22, 2019 · [Twitter for Android](#)

NON-EMERGENCY



Drain Doctor B'ham
@DrainDoctorBham

At Drain Doctor [#Birmingham](#) [#Plumbing](#) and Drainage we're proud to provide an instant response to emergency plumbing call-outs from commercial customers, with zero/minimum business downtime (night-time working at no additional price to avoid disruption). [#Drains](#)
[#Emergency](#)



KEY TAKEAWAYS / RECOMMENDATIONS

Key Takeaway 1

Logistic regression performed better overall for our classification model. Although both scored well, random forest was overfit to our data

Key Takeaway 2

The effects of certain keywords and how they are categorized affects the accuracy of our data (i.e. "sharknado"). Creating a lexicon of words to train will help us to continually tune and find relevant emergencies

Recommendation 1

Always create a verification process where someone reviews the tweets for legitimate cause for action. Although the score is fairly accurate we want to ensure the right precautions are taken

Recommendation 2

Continued research within the Twitter API to generate more accurate locations as well as number of emergency tweets coming from that location will give more meaning to the received alert

A photograph of a man from behind, wearing a dark blue FEMA vest over a tan shirt, a dark baseball cap, and glasses. He is looking towards a damaged roof structure with exposed wooden beams. The image is framed by a light blue border with orange and purple geometric shapes.

THANK YOU

Does anyone have any questions?

Sidd Nirmal
Krista Johnson

SK
RESEARCH