University at Buffalo
School of Engineering and Applied Sciences

Using Social Media for Search and Rescue: Case Study on Hurricane Harvey

Abhinav Khare, Dr. Qing He, Dr. Rajan Batta
Department of Industrial and Systems Engineering, University at Buffalo

Outline

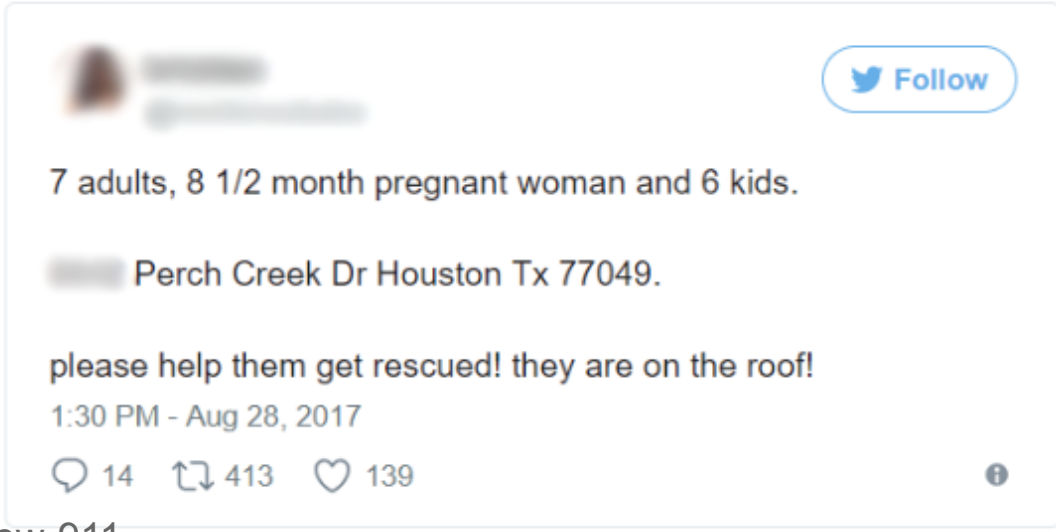
- Motivation
- Objective
- Data
- Data Filtering
- Methodology Overview
- Results
- Future work

Motivation

- **Time Magazine:** 'Please send help' Hurricane Harvey Victims Turn to Twitter and Facebook.
http://time.com/4921961/hurricane-harvey-twitter-facebook-social-media/

- **Fox News**: Tropical Storm Harvey : Is Twitter become the new 911?
https://www.foxnews.com/tech/tropical-storm-harvey-is-twitter-becoming-the-new-911

- **Npr.org**: Facebook and Twitter replace 911 calls for stranded in Houston.
https://www.npr.org/sections/alltechconsidered/2017/08/28/546831780/texas-police-and-residents-turn-to-social-media-to-communicate-amid-harvey



Ref : Z.Yang, L.Hoang Nguyen, J. Stuve, G. Cao, F.Jin Harvey Flooding Rescue in Social Media. 2017 IEEE International Conference on Big Data (BIGDATA)

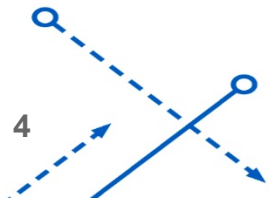
Objective

- Build a methodology that utilizes social media information in search and rescue models.

- There are two sub-objectives to reach this objective:

1. Sub-Objective 1 : Extracting relevant social media information from the complex social media data. (Build a tweet Classifier)

2. Sub-Objective 2 : Building a search and rescue model based on this information. (Queuing model)



Data

We accessed 36,139,367 tweets with the following 41 columns that include from Texas :

- TWEET_ID
- TWEET_TEXT
- USER_ID
- USER_NAME
- DATE
- TIME
- TIMEZONE
- HASHTAGS
- LATITUDE
- LONGITUDE
- BOUNDING BOX

Date Filtering

Using Dates

- 27/08/2017 - 31/08/2017

Using Keywords (from Regular Expressions)

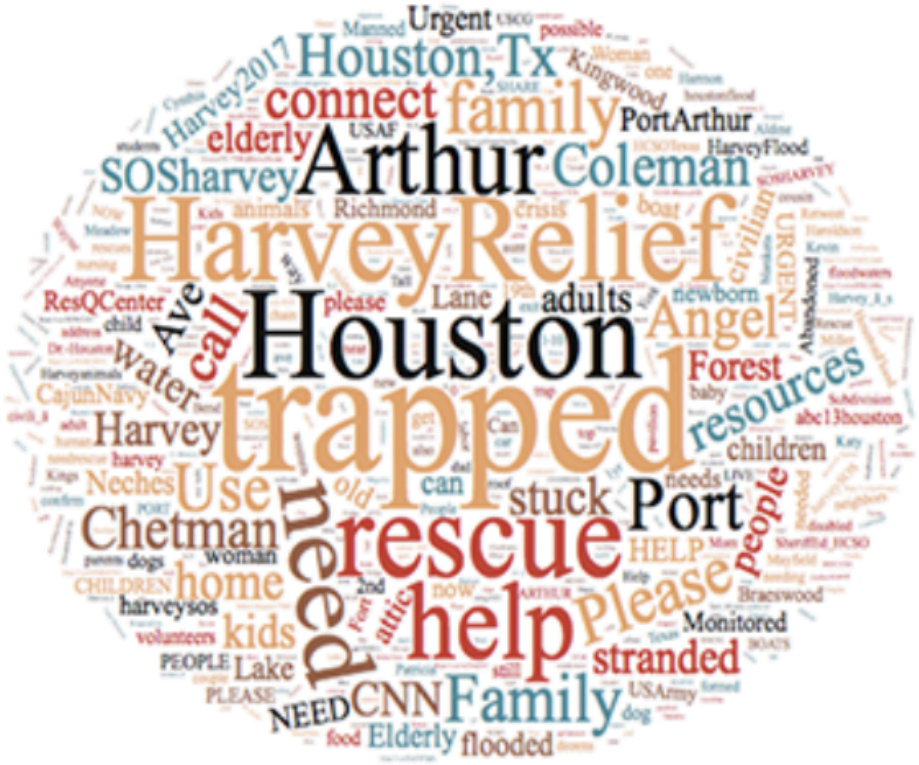
- Help
- Rescue
- SOS
- Trapped
- Urgent
- Stuck
- Stranded
- Please
- Need
- HarveyRescue

Filtered Data

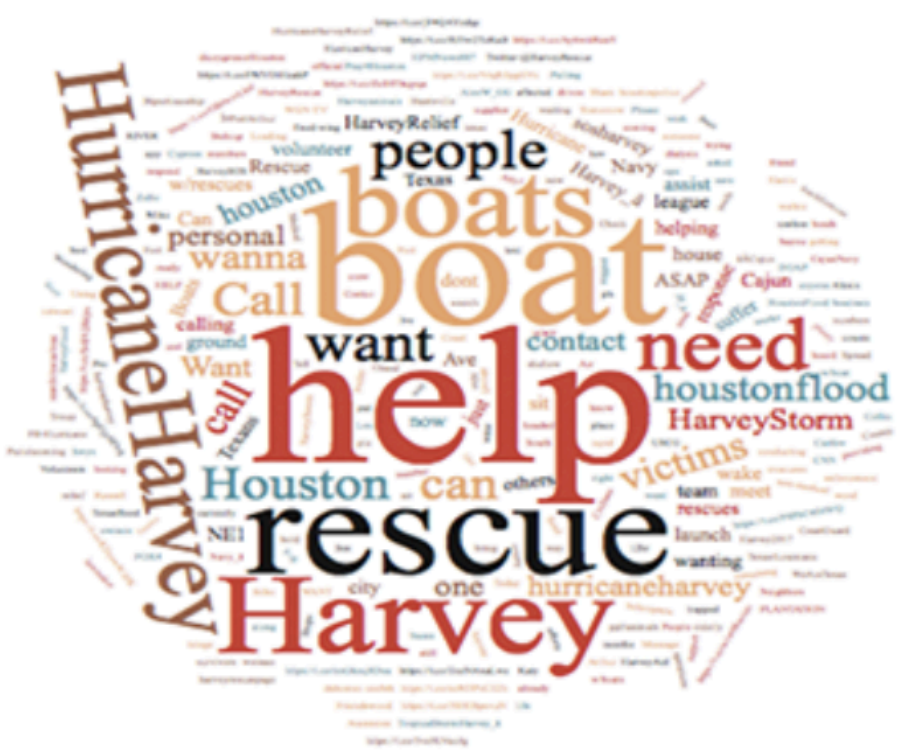
23,455 tweets of the following kind of tweets were filtered out :

TWEET TEXT	REQUEST/ VOLUNTEER
SOS: Day 2 of being trapped inside my neighborhood, and many days to go. Will offer $100 for a can of dip https://t.co/4HX534nxCK	Request
#HurricaneHarvey: Anyone in #Charlestowncolony area help is needed ASAP!!! #HTX #sos‰Û_ https://t.co/gw1Qluq67s	Request
Orange, Texas needs help!!! #HarveySOS	Request
My friend needs help please 5726 Charles ave elderly woman in there!! #harveysos	Request
I hope everyone is okay if anyone needs help Dm me	Volunteer
@SOS_DISPATCH I'm able and willing to help! If someone with a boat can pick me up in league city! Thanks	Volunteer
DM ME IF YOU NEED HELP IN KATY TX!! #harveysos #HarveyRescue #HoustonFlood	Volunteer
We don't have enough help out here man please God send someone #PortArthur #harveySOS #HarveyFlood	Volunteer

Previous Results in Literature on Sub-objective 1



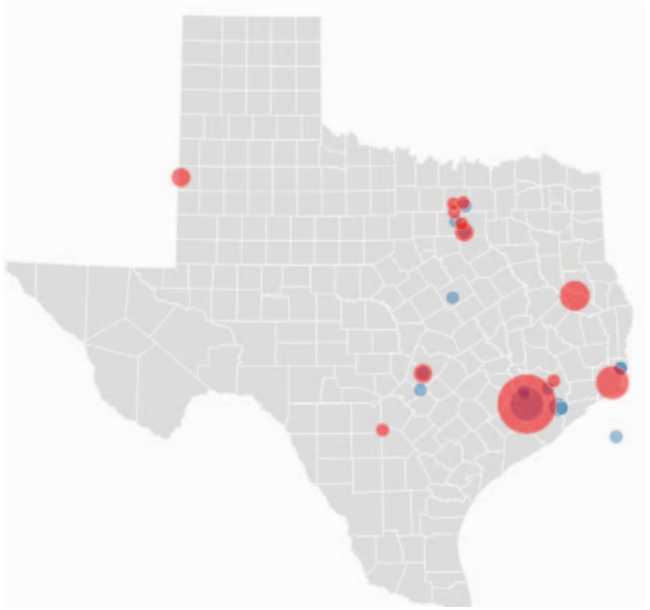
a) Word cloud for request tweets

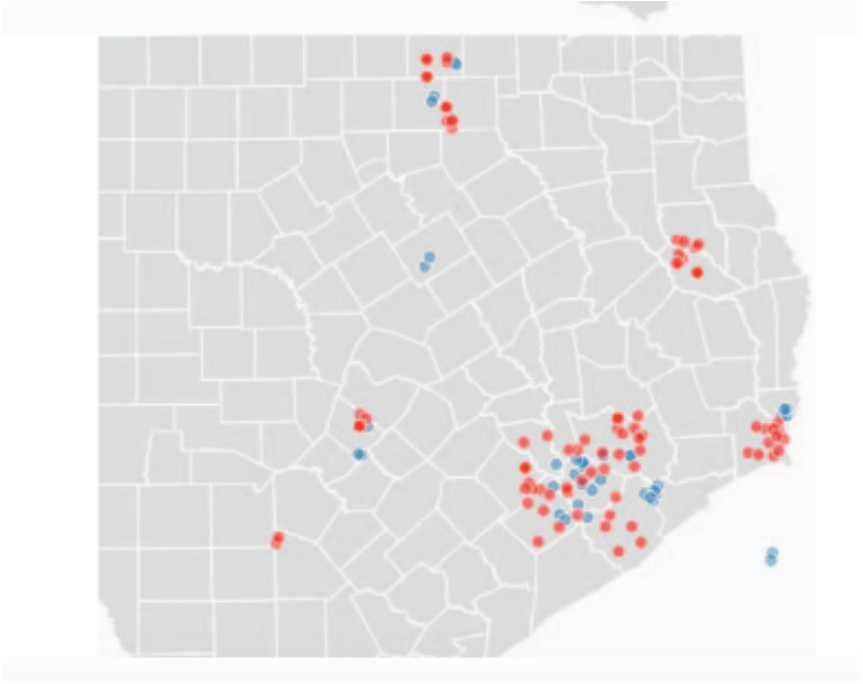


b) Word cloud for volunteer tweets

Ref : Z.Yang, L.Hoang Nguyen, J. Stuve, G. Cao, F.Jin Harvey Flooding Rescue in Social Media. 2017 IEEE International Conference on Big Data (BIGDATA)

Previous Results in Literature on Sub-objective 1





Distribution of victims needing help (red dots) and volunteers with boats (blue dots) in Texas and Houston respectively

Ref : Z.Yang, L.Hoang Nguyen, J. Stuve, G. Cao, F.Jin Harvey Flooding Rescue in Social Media. 2017 IEEE International Conference on Big Data (BIGDATA)

Previous Results in Literature on Sub-objective 1

Previous Tweet Classifiers for search and rescue tweets

	Precision	Recall	F measure	Accuracy
Log. Regression	0.248	0.658	0.360	0.703
KNN	1.000	0.413	0.584	0.926
CART	0.710	0.579	0.638 0	0.917
SVM 2	0.606	0.793	0.687	0.930

Ref : Z.Yang, L.Hoang Nguyen, J. Stuve, G. Cao, F.Jin Harvey Flooding Rescue in Social Media. 2017 IEEE International Conference on Big Data (BIGDATA)

Tweet Classifier

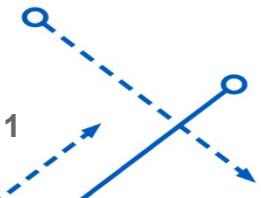
- Developed a two tweet classifiers of our own to identify tweets for rescue and volunteering.

- Tweet Classifier 1

1. Bag of words model with an SVM classifier.
2. In bag of words model, M important words identified from the tweets and each tweet is one hot-encoded vector of size M.

- Tweet Classifier 2

1. LSTM classifier in which words are input as word embedding's obtained using a Word2Vec Skip-gram model.
2. In Word2vec model the corpus of text produces a vector space, typically of several hundred dimensions, with each unique word in the corpus being assigned a corresponding vector in the space.

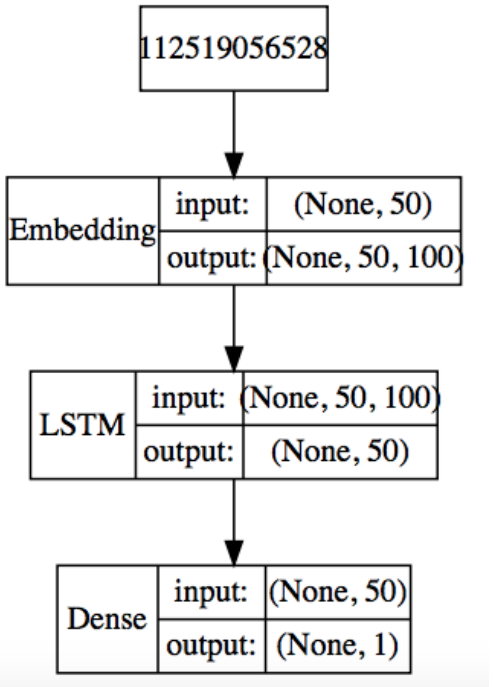


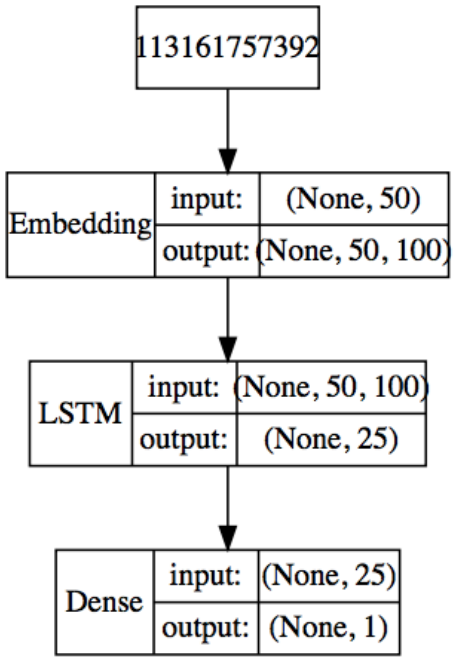
Tweet Classifier

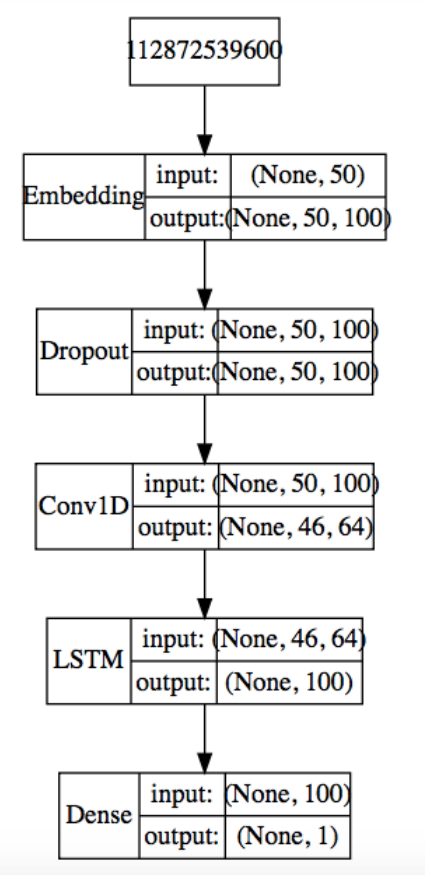
- For both the Tweet Classifiers, a training, validation and test set was created.
- Training set contained 2000 tweets labeled manually.
- Validation set contained 500 tweets for model selection.
- Test set contained 500 tweets to test the selected model.

Tweet Classifier 2 (RNN)

- Multiple architectures of RNN were tried and tested on the validation set.
- Hyper-parameters like dropout ratio and early stopping were also varied
- Following were the best 3 architecture. The last one also adds a layer of convolution before the LSTM layer.
- The first classifier with 50 units in the LSTM layer performed the best on the validation set







Results of Tweet Classifer

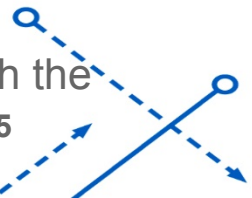
	Precision	Recall	F measure	Accuracy
Tweet Classifier 1 (SVM)	0.9831247	0.7612366	0.8580682	0.961
Tweet Classifier 2 (LSTM)	0.9852071	0.7894737	**0.8765465**	0.965
Log. Regression	0.248	0.658	0.360	0.703
KNN	1.000	0.413	0.584	0.926
CART	0.710	0.579	0.638 0	0.917
SVM	0.606	0.793	0.687	0.930

- These are the results of the accuracy measures on the validation set .
- Both the classifier outperform the classifiers of Yang et al., LSTM being the best classifier.

Literature on Subobjective 2: Spatial Queue Model

- Our sub-objective was to build a spatial queuing model for search and rescue.
- In spatial queue of search and rescue the servers are the volunteers and customers are the calls for rescue.
- Existing Spatial Queueing Models for searcha and rescue in Literature

1. FCFS M/M/c queuing model
2. Priority M/M/c, queuing model : tweets with words like *'grandma, grandpa, senior, old, baby, kid, pregnant, sick, ill, dangerous'* are given priority
3. Hybrid model
   - The hybrid scheduling policy is non-preemptive.
   - If the server is idle, incoming requests will be scheduled based on FCFS.
   - The priority value is assigned to every request upon its creation, and it can be changed due to an emergency situation.
   - When a server is available, requests with the highest priority will be scheduled first. If all the requests are with the same priority, then the request with the earliest finish time will be scheduled first.



Sub-objective 2 – Queuing Model

- In our Spatial queue we take volunteer management into account

- **Volunteer management** in disaster management become important due to two reasons:
  1. Large number of volunteers.
  2. Unreliable Volunteers

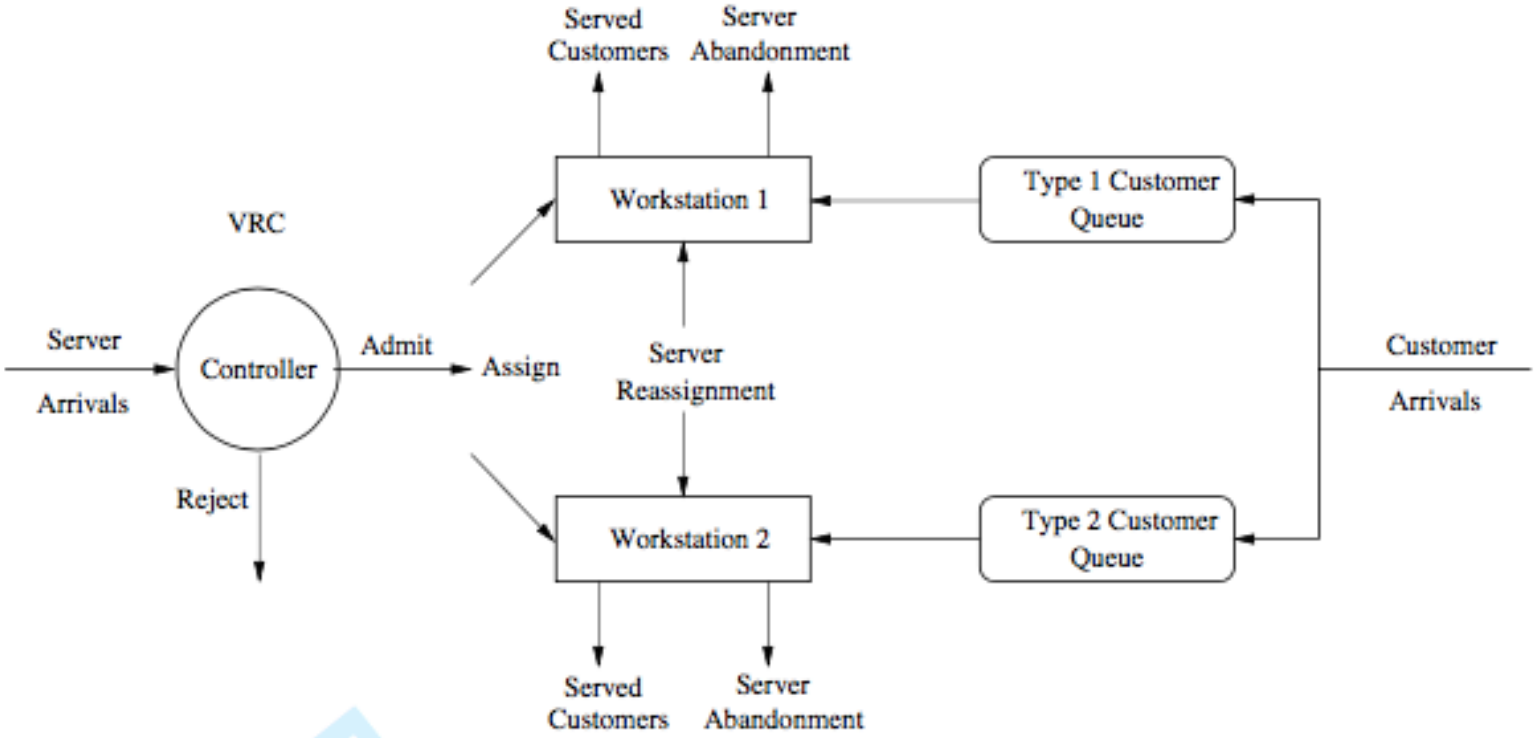
People line up to volunteer CREDIT: MICHAEL CIAGLO /HOUSTON CHRONICLE VIA AP

Volunteers listen to instructions CREDIT: MICHAEL CIAGLO/HOUSTON CHRONICLE VIA AP

Sub-objective 2 – Queuing Model

In this scenario, the spatial queues set up should have two properties:

1. They should model volunteer reliability (server abandonment rate)
2. They should be complemented with a decision making model to select volunteers



Judging the reliability of volunteers

Approach 1 : Reliability score based on the tweet

Drawbacks :

1. How to choose the scale for the score ?
2. The score is dependent on the scorer's judgment.
3. No ground truth about the person's credibility/reliability.
4. Cannot manually score thousands of tweet.

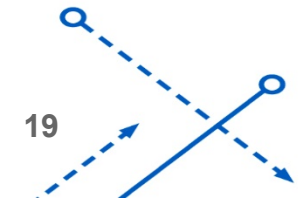
TWEET_TEXT	Reliability score
Who has a truck is willing to go volunteer ? I'm thinking about asking my dad to take his boat down to Houston and helping out	2
You got a boat for me to get there? https://t.co/hoAraaLKpp	0
MY DAD IS OUT ON HIS BOAT IN SOUTH HOUSTON LET ME KNOW IF ANYONE NEEDS RESCUINGâ€¼ï¸	3
Man I wanna help with some rescues but I don't have boat and my truck ain't lifted high enough ðŸ˜’	1

Future Work

- Approach 2: Truthfulness score for the tweeters using his/her previous tweets (accessed via REST API)

- Approach 3 : Natural Language Inference Models .

Premises	Label	Hypothesis
You got a boat for me to get there?	Contradiction	This is a reliable volunteer
Man I wanna help with some rescues but I don't have boat and my truck ain't lifted high enough	Neutral	This is a reliable volunteer
MY DAD IS OUT ON HIS BOAT IN SOUTH HOUSTON LET ME KNOW IF ANYONE NEEDS RESCUINGâ€¼ï	Entailment	This is a reliable volunteer



Thank you