Leveraging News and Media for Situational Awareness

Sara Almeida, Callie Bouzane, Marielle Marcus

The Problem

When an emergency or disaster happens, the affected community wants to get as much information about it as possible. But this information is not compiled in one central location.

Our Solution

A **content-based recommender** system:

- Select an article about the emergency at hand

- Pull up similar articles to get comprehensive information all in one place

- Centralized data center where the public can learn more about emergencies of interest

High Level Process

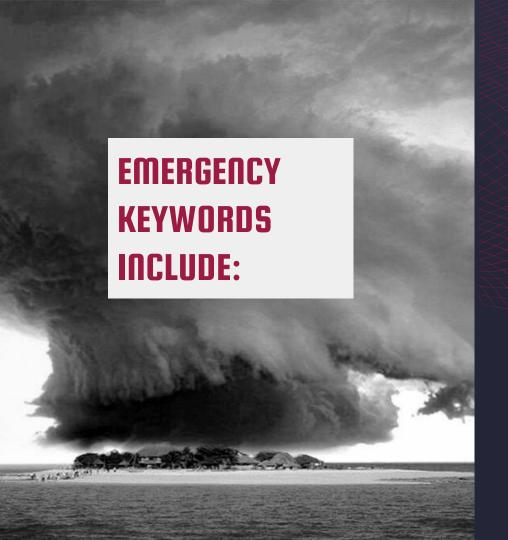| Gather Data (News API) | → | Parse Data (RegEx & NLP) | → | Build Recommender System |

News API

30,000 news sources

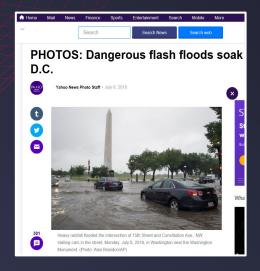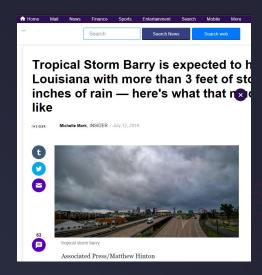nº news media

3,400

articles scraped

EMERGENCY KEYWORDS INCLUDE:

1. Flood

2. Heat wave

3. Blizzard

4. Tropical Storm

5. Earthquake

6. Fires

7. Hurricane

8. Polar Vortex

9. Tornado

Examples of articles

PHOTOS: Dangerous flash floods soak D.C.

Yahoo News Photo Staff · July 8, 2019

Heavy rainfall flooded the intersection of 15th Street and Constitution Ave., NW stalling cars in the street, Monday, July 8, 2019, in Washington near the Washington Monument. (Photo: Alex Brandon/AP)

301

Tropical Storm Barry is expected to h Louisiana with more than 3 feet of sto inches of rain — here's what that me like

INSIDER   Michelle Mark, INSIDER · July 12, 2019

tropical storm barry

Associated Press/Matthew Hinton

63

★ StarTribune

LOG IN

ALL SECTIONS

Home   Local   Sports   Business   Opinion   Variety   Obituaries   Classifieds   A

MORE FROM THE HOMEPAGE

...uses in Chippewa | State lawmakers agree on insulin program, but funding in question | Klobuchar's Senate the partisan divide

LOCAL

Tornadoes touch down in northern Minnesota

Stormy weather forecast for the Twin Cities metro area failed to materialize so far.

By Zoë Jackson Star Tribune | JULY 17, 2019 — 11:24PM

SEVERE THREAT WEDNESDAY NIGHT

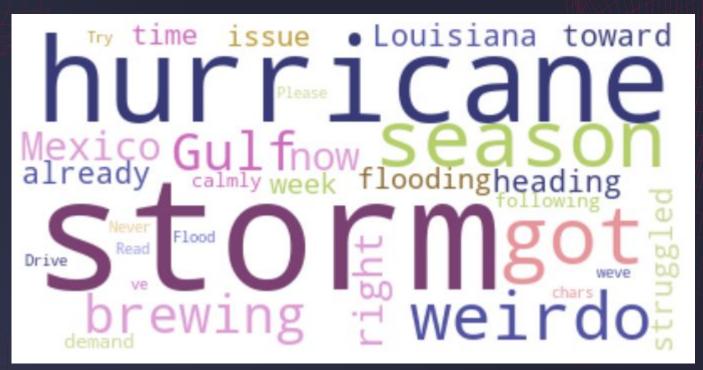▶ Video (01:05) : Evening Star Tribune Local Twin Cities (Minneapolis/St. Paul)

Data Preprocessing

Data = news (Title + Description + Article Content)
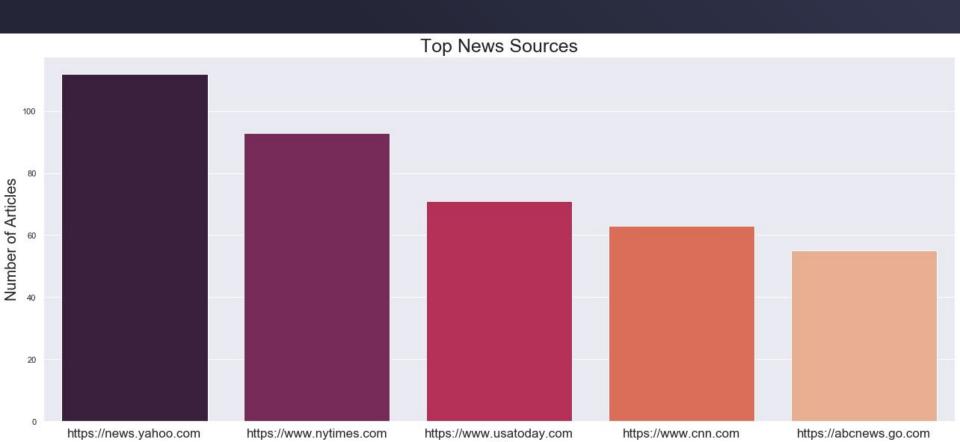
Data Preprocessing

TFIDF
- Used to identify the distinguishing features of documents relative to others
- Hyperparameters:
  - Maximum number of features
  - Maximum & minimum occurrences among documents
  - N-Gram Range
  - Stopwords removed

Source Extraction

Top News Sources

Recommender in Action

Steps:
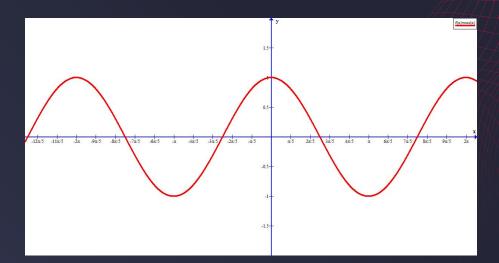1. Search for a term & [optional] location
2. Choose an article you would like to learn more about
3. Search the article title in our recommender
4. Up to 10 articles will be displayed, with similarity scores of at least 0.4

```
get_recommendations('Torrential rain floods Washington area')
```

	Article Title	Similarity Score	Article URL
0	PHOTOS: Dangerous flash floods soak Washington, D.C.	0.622848	https://news.yahoo.com/photos-dangerous-flash-floods-soak-washington-dc-170414524.html
1	Thunderstorms in Washington DC have caused flooding in the White House, road washouts, and indoor waterfalls	0.501093	https://www.insider.com/flooding-in-washington-dc-caused-road-washouts-and-indoor-waterfalls-2019-7
2	Flooding reported on several roadways amid torrential downpour - CP24 Toronto's Breaking News	0.478079	https://www.cp24.com/firefighters-responded-to-12-water-rescue-calls-amid-torrential-downpour-1.4511514
3	Flash-Flood Watch Threatens Washington's July 4 Celebrations	0.415850	https://www.bloomberg.com/news/articles/2019-07-04/washington-at-risk-of-flash-flooding-as-alert-issued-for-july-4
4	Nearly 5 million people under flash flood watch	0.401775	https://www.cnn.com/videos/weather/2019/07/13/barry-louisiana-new-orleans-weather-saturday-vpx.cnn

Recommender in Action

Evaluation Metric: Cosine Similarity
- Ranges from -1 to 1
- Score closer to 1 represents a smaller "distance" between the two vectors (representing individual documents)

Next Steps

- Add more news sources
- Upgrade to get more real-time information
- Introduce an interface where users can search in a Google-esque search engine to generate relevant articles

Thank You!