CS 106S
Sentiment Analysis

materials:
bit.ly/cs106s-download-nlp

announcements

- email list has been created
- reminder to fill out checkoff forms
- office hours are mondays 9-10 pm in 200-015
- preliminary syllabus released
- apply to cs+sg fellowships (due tonight)

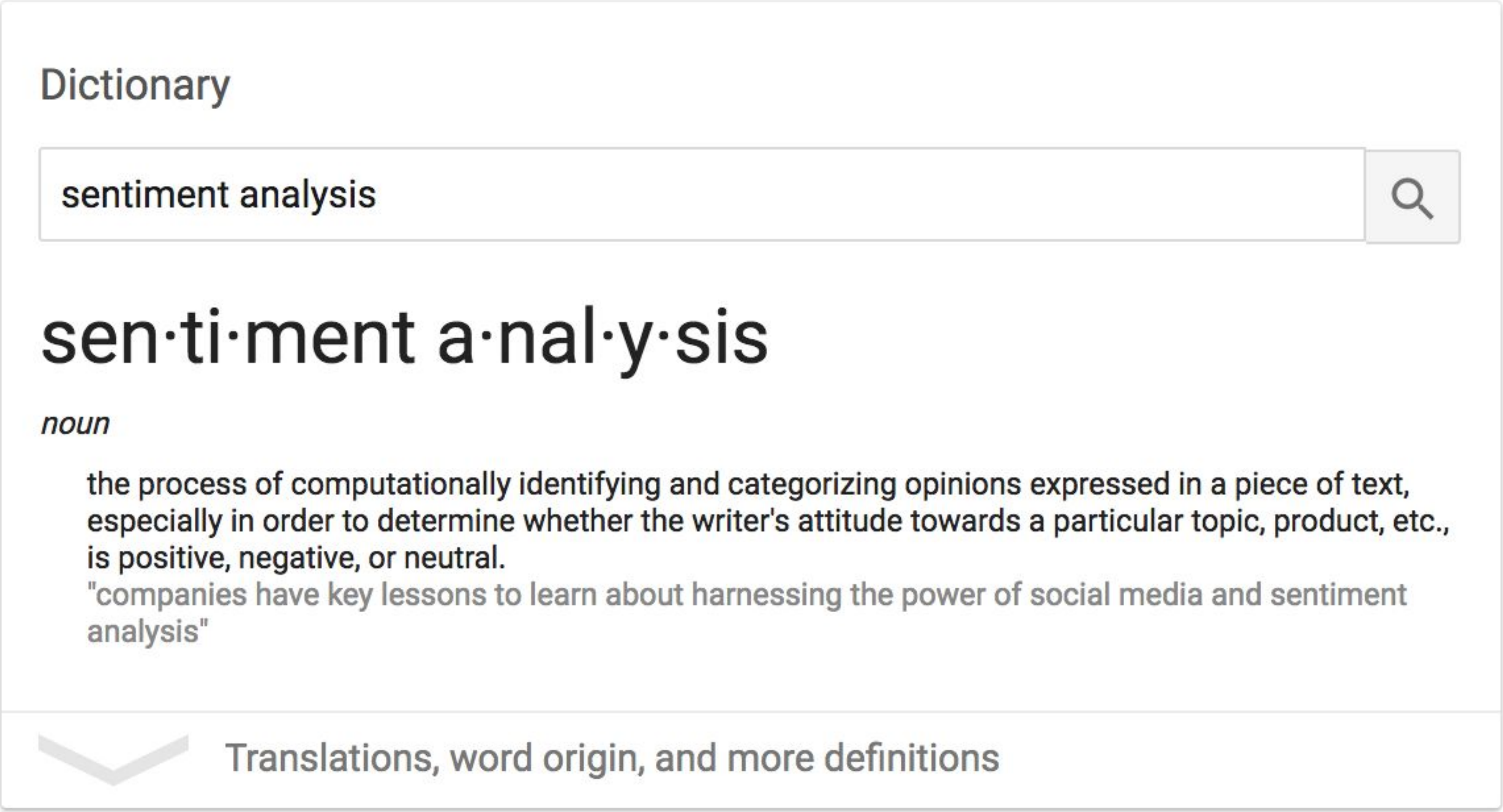
overview

1. sentiment analysis overview
2. the problem:
   refugee sentiment on twitter
3. our solution
4. assignment

attendance checkoff

bit.ly/cs106s-checkoff-nlp

sentiment analysis



Dictionary

sentiment analysis

sen·ti·ment a·nal·y·sis

noun

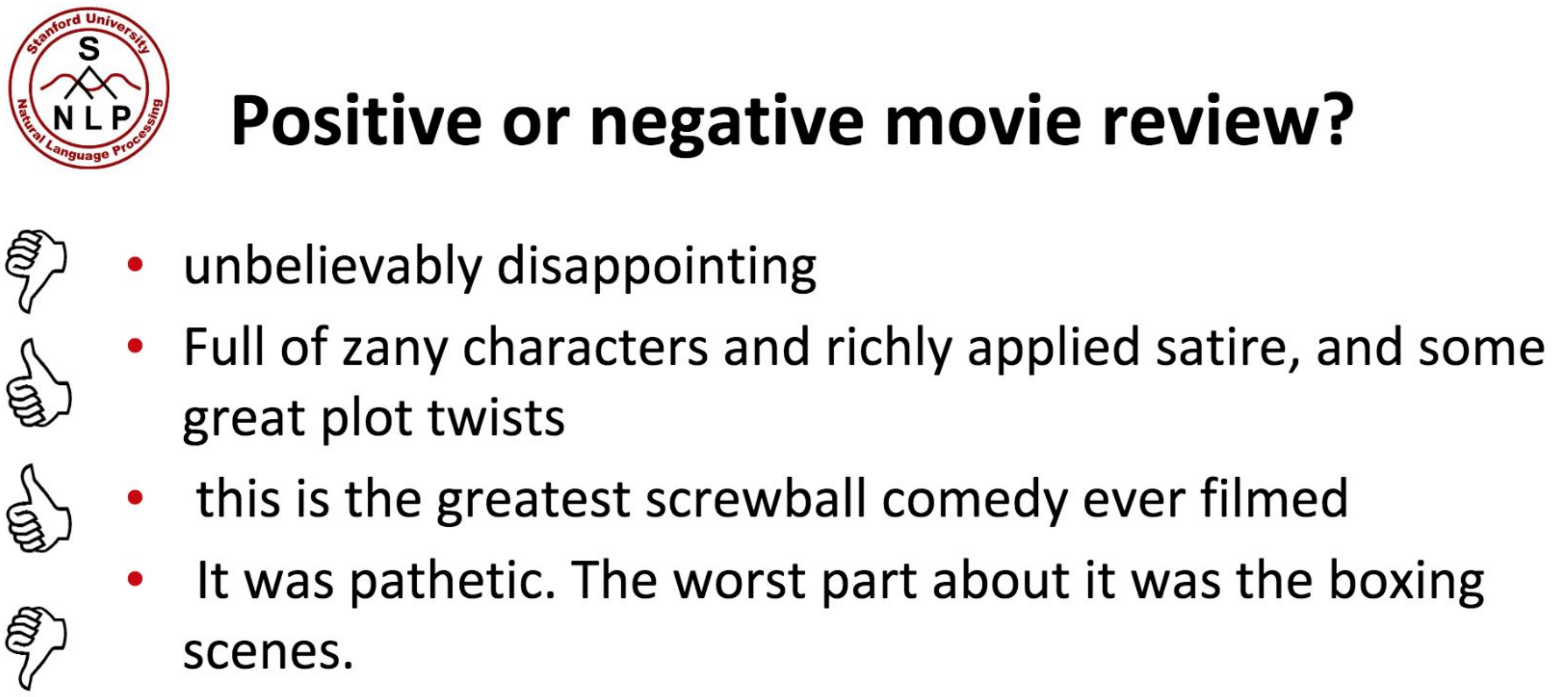
the process of computationally identifying and categorizing opinions expressed in a piece of text, especially in order to determine whether the writer's attitude towards a particular topic, product, etc., is positive, negative, or neutral.
"companies have key lessons to learn about harnessing the power of social media and sentiment analysis"

Translations, word origin, and more definitions

how?

- look at a piece of text
- extrapolate “features” in the text
- common features:
  - is a **word** positive or negative?
  - is a **phrase** positive or negative?
- classify the piece as positive or negative

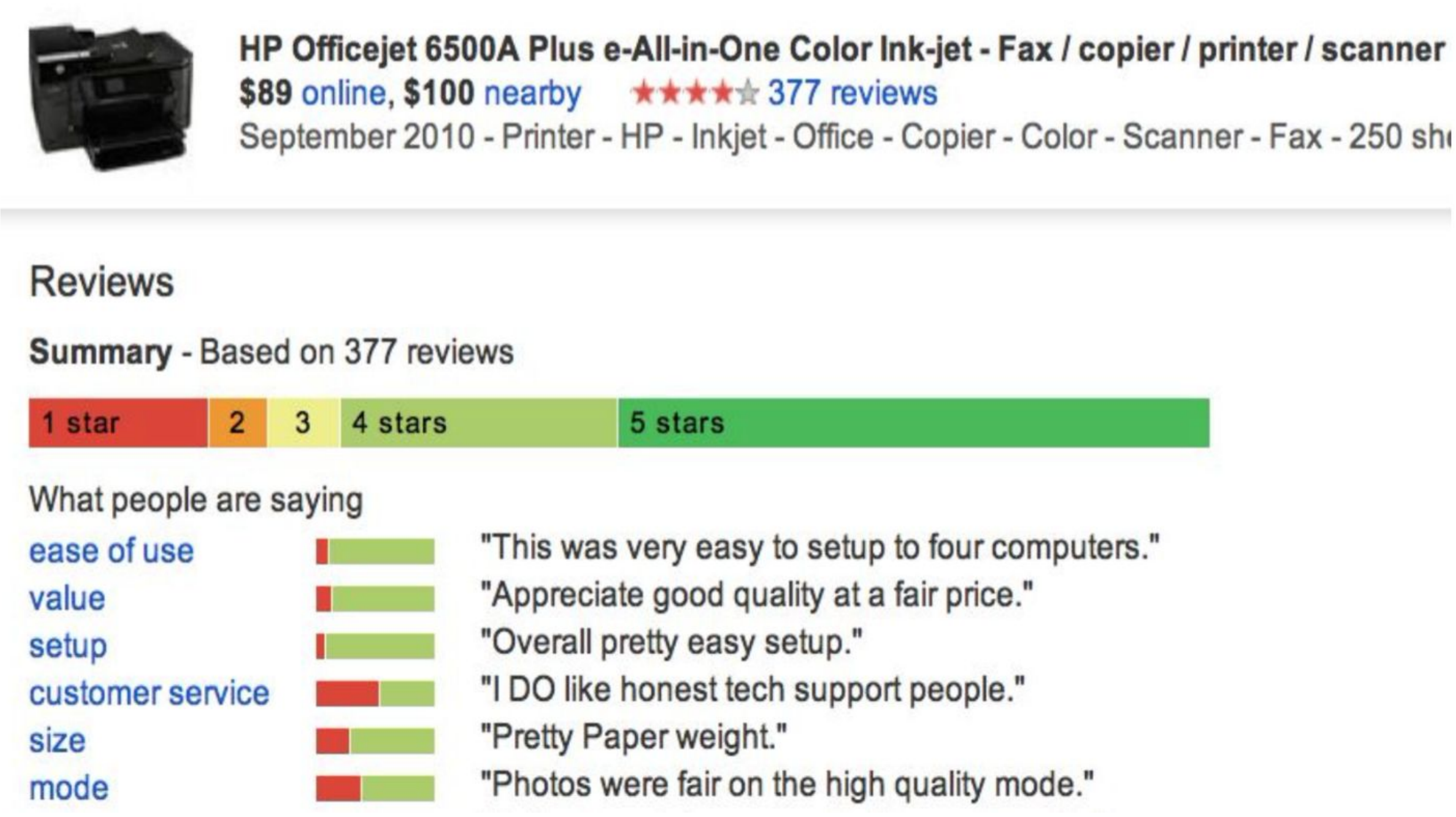
uses



Positive or negative movie review?

- 👎 unbelievably disappointing
- 👍 Full of zany characters and richly applied satire, and some great plot twists
- 👍 this is the greatest screwball comedy ever filmed
- 👎 It was pathetic. The worst part about it was the boxing scenes.

Google Product Search



HP Officejet 6500A Plus e-All-in-One Color Ink-jet - Fax / copier / printer / scanner

$89 online, **$100** nearby ★★★★☆ 377 reviews

September 2010 - Printer - HP - Inkjet - Office - Copier - Color - Scanner - Fax - 250 sh

Reviews

Summary - Based on 377 reviews

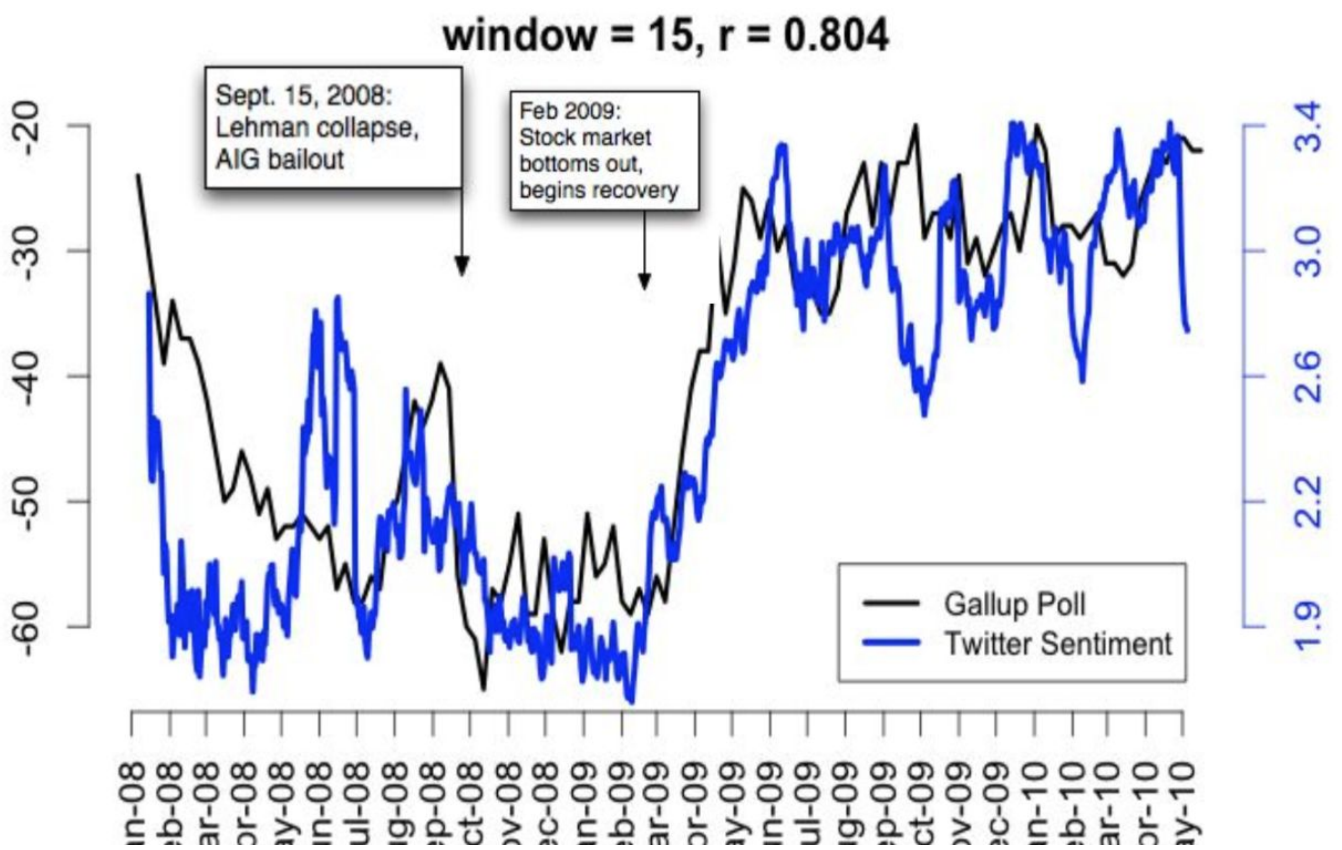
1 star	2	3	4 stars	5 stars

What people are saying

ease of use	"This was very easy to setup to four computers."
value	"Appreciate good quality at a fair price."
setup	"Overall pretty easy setup."
customer service	"I DO like honest tech support people."
size	"Pretty Paper weight."
mode	"Photos were fair on the high quality mode."

Twitter sentiment versus Gallup Poll of Consumer Confidence

Brendan O'Connor, Ramnath Balasubramanyan, Bryan R. Routledge, and Noah A. Smith. 2010. From Tweets to Polls: Linking Text Sentiment to Public Opinion Time Series. In ICWSM-2010



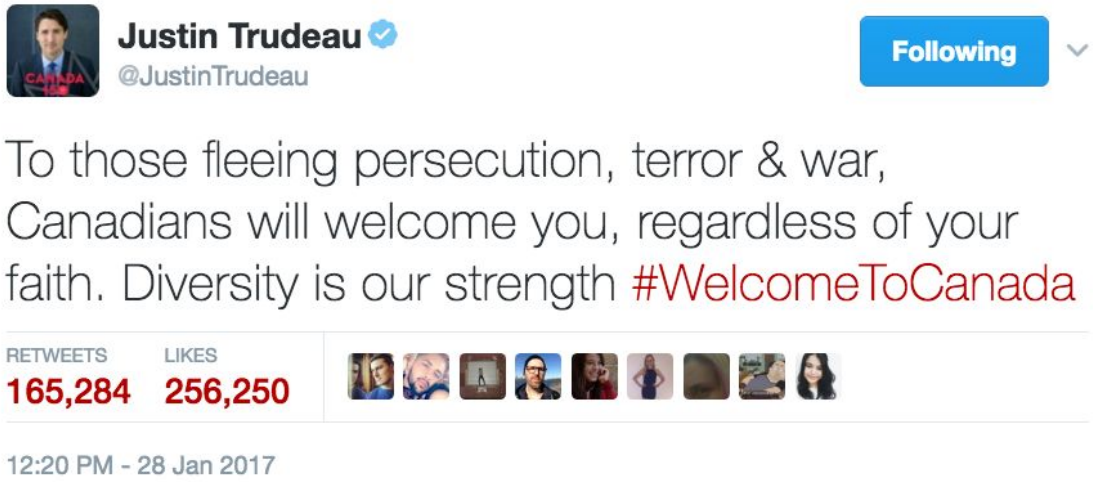
uses

- movie reviews: what does the public think of a movie?
- products: what do people think of the new iPhone?
- public sentiment: reaction to news? current events?
- politics: what do people think of x issue?
- prediction: what will the outcome of an election be based on social media data?

uses

- movie reviews: what does the public think of a movie?
- products: what do people think of the new iPhone?
- public sentiment: reaction to news? current events?
- politics: what do people think of x issue?
- prediction: what will the outcome of an election be based on social media data?

the issue: understanding refugee sentiment on twitter

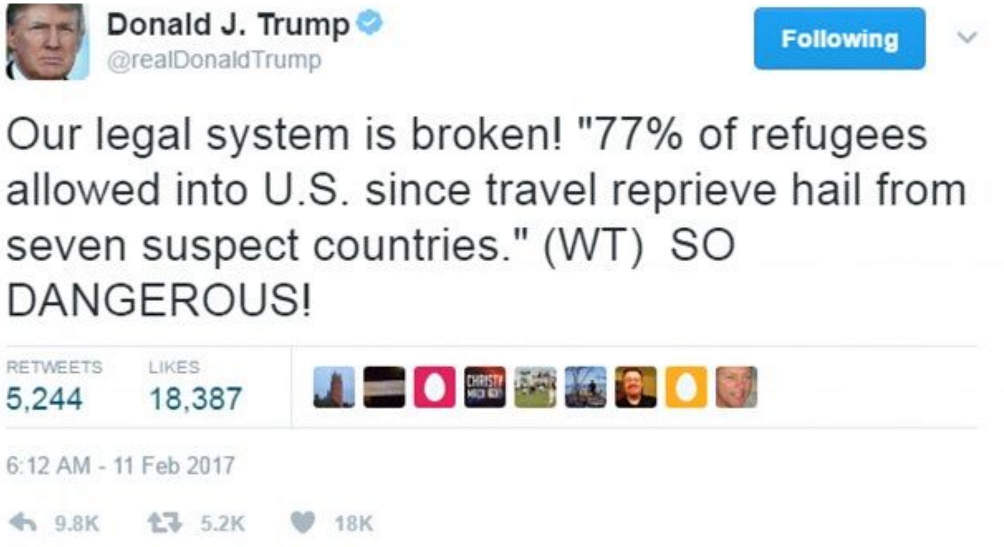


Justin Trudeau
@JustinTrudeau

To those fleeing persecution, terror & war, Canadians will welcome you, regardless of your faith. Diversity is our strength #WelcomeToCanada

RETWEETS 165,284 LIKES 256,250

12:20 PM - 28 Jan 2017



Donald J. Trump
@realDonaldTrump

Our legal system is broken! "77% of refugees allowed into U.S. since travel reprieve hail from seven suspect countries." (WT) SO DANGEROUS!

RETWEETS 5,244 LIKES 18,387

6:12 AM - 11 Feb 2017

9.8K 5.2K 18K

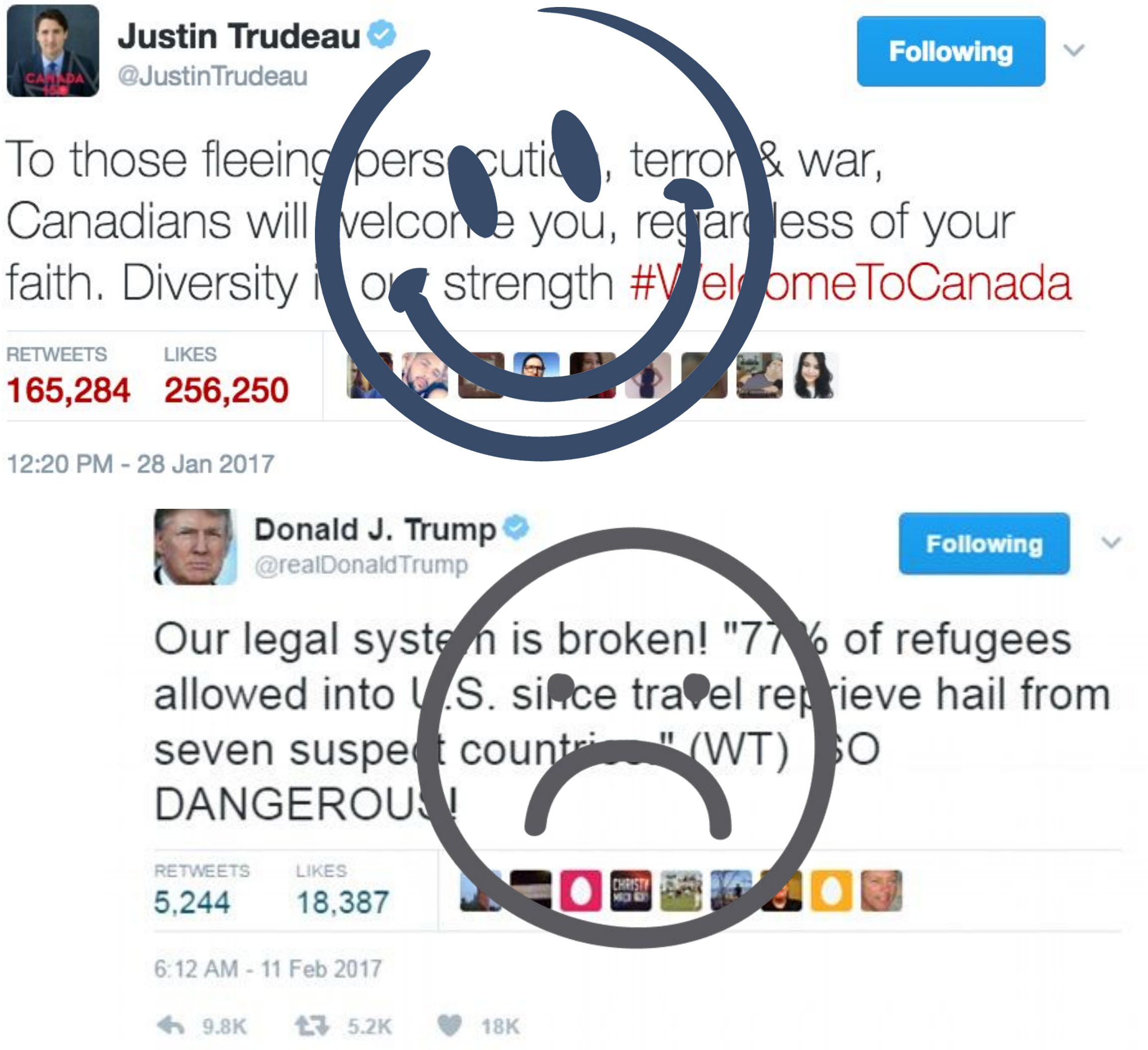
can we accurately classify tweets as anti-refugee or pro-refugee?

why?

- bots that detect hateful tweets and respond with educational resources
  - sway public opinion
- metrics on demographics of anti-refugee actors to inform resistance and education efforts
- understanding general opinion in america
  - spring '17 cs50 project with the united nations!

mid-class q&a!

our solution: build a refugee sentiment tweet classifier



Justin Trudeau
@JustinTrudeau

To those fleeing persecution, terror & war, Canadians will welcome you, regardless of your faith. Diversity is our strength #WelcomeToCanada

RETWEETS 165,284 LIKES 256,250

12:20 PM - 28 Jan 2017

Donald J. Trump
@realDonaldTrump

Our legal system is broken! "77% of refugees allowed into U.S. since travel reprieve hail from seven suspect countries." (WT) SO DANGEROUS!

RETWEETS 5,244 LIKES 18,387

6:12 AM - 11 Feb 2017

key approach

- give each word in a tweet a value
- represents how pro/anti-refugee sentiment it is

```
var wordMap = {
    "illegal": -3,
    "welcome": 2,
    "fear": -1,
    "innocent": 4,
    ...
}
```

creating wordMap

1. iterate through each training tweet
2. get tweet's classifictaion
3. for each word in the tweet
   a. stem the word
   b. if tweet is anti-refugee, -1 to word's score in wordMap
   c. else +1

creating wordMap

1. iterate through each training tweet
2. get tweet's classification
3. for each word in the tweet
    a. stem the word
    b. if tweet is anti-refugee, -1 to word's score in wordMap
    c. else +1

Example: "We love refugees #refugees" - Positive Tweet

creating wordMap

1. iterate through each training tweet
2. get tweet's classification
3. for each word in the tweet
    a. stem the word
    b. if tweet is anti-refugee, -1 to word's score in wordMap
    c. else +1

Example: "We love refugees #refugees" - Positive Tweet

creating wordMap

1. iterate through each training tweet
2. get tweet’s classification
3. for each word in the tweet
    1. stem the word
    2. if tweet is anti-refugee, -1 to word’s score in wordMap
    3. else +1

Example: “We love refugees #refugees” - Positive Tweet

creating wordMap

1. iterate through each training tweet
2. get tweet's classification
3. for each word in the tweet
   a. stem the word
   b. if tweet is anti-refugee, -1 to word's score in wordMap
   c. else +1

Example: "We love refugees #refugees" - Positive Tweet

creating wordMap

1. iterate through each training tweet
2. get tweet's classification
3. for each word in the tweet
    a. stem the word //var stemmedWord = stemmer(word);
    b. if tweet is anti-refugee, -1 to word's score in wordMap
    c. else +1

Example: "We love refugees #refugees" - Positive Tweet

creating wordMap

1. iterate through each training tweet
2. get tweet’s classification
3. for each word in the tweet
    a. stem the word //var stemmedWord = stemmer(word);
    b. if tweet is anti-refugee, -1 to word’s score in wordMap
    c. else +1

+1 +1 +1 +1

Example: “We love refugees #refugees” - Positive Tweet

wordMap

```
var wordMap = {
    “we”: 1,
    “love”: 1,
    “refugees”: 1,
    “#refugees”: 1,
    …
}
```

tip #1: tweet structure

For each tweet, use its **text** and **classification**.

```
For (var i = 0; i < trainTweet.length; i++) {
    var tweet = trainTweet[i];
    var tweetText = tweet.tweet;
    var tweetTextLower = tweet.tweet.toLowerCase();
    var tweetLabel = getScore(tweet.classification)
}
```

tip #2: getScore() is helpful

Returns -1 if pro-refugee, -1 if anti-refugee.

```
For (var i = 0; i < trainTweet.length; i++) {
    var tweet = trainTweet[i];
    var tweetText = tweet.tweet;
    var tweetTextLower = tweet.tweet.toLowerCase();
    var tweetLabel = getScore(tweet.classification)
}
```

tip #3: make the tweet lower case!

Why?

```
For (var i = 0; i < trainTweet.length; i++) {
    var tweet = trainTweet[i];
    var tweetText = tweet.tweet;
    var tweetTextLower = tweet.tweet.toLowerCase();
    var tweetLabel = getScore(tweet.classification)
}
```

classifying test tweets

1. iterate through each test tweet in testTweets
2. initialize a variable to store our “score” of the tweet.
3. for each word in the tweet:
    a. stem the word*
    b. if word is in wordMap, add its score to our variable
4. once we go through all words in a tweet, classify!
    a. if score is > 0, classify as pro-refugee (false)
    b. if score is < 0, classify as anti-refugee (true)

```
var wordMap = {
    "hate": -5,
    "love": +4,
    "all": +1,
    "refugees": +1,
    ...
}
```

"Don't hate. Love all refugees #refugees" – Positive

score = -5 + 4 + 1 + 1 +1 = +2

category = positive

set myGuesses[testTweet.tweetID] = false

// if this were a anti-refugee tweet, we set to true and that’s the basic model!

pitfalls

what doesn’t this model account for?

improving our performance

tf-idf

(term frequency - inverse document frequency)

tf-idf

- term frequency
  - number of times t appears in document/total terms
- inverse document frequency
  - log_e(total documents/documents with term t)
- tf-idf = tf * idf
- want words that are important in classifications, not across documents

intuition

- term frequency
  - insteading of adding +1 or -1 for each word, add +1/n or -1/n where n is number of words in the tweet
- inverse document frequency
  - maintain separate mapping of word to number of documents it’s in
- join these two mappings at the end (iterate through both and multiply term frequency by the inverse of the document frequency)

assignment 1 - nlp

- finish today's assignment (including tf-idf)
- submit through email to cs106s-spr1718-staff@lists.stanford.edu
- **due: friday, april 20th at noon**

next week...

compas

an introduction to algorithmic bias

See you next time, friends!