

Classifying Company & Product Sentiment Based on Tweets at a Tech Conference

MARK P. & TIM M. JANUARY 2021

MT Head Consulting

- Customer feedback insights
- Product and service ideas
- Product positioning
- Marketing & PR planning



“Keeping an ear to the ground, and an open mind...”



2011 Austin, TX

@ SXSW Interactive:

- iPad 2 launch
- Phone apps still wonky
- Google Maps sends you to Death Valley
- Earthquake / Tsunami in JP



Can we glean useful insights from conference tweets?

- Tweets are frequent & ubiquitous
- Tech conference with focus on companies, products, services
- **Positive** tweets: what resonates
- **Negative** tweets: issues & complaints

Goal: Accurately classify tweets by sentiment



Just when you thought social couldn't get more overblown at #sxsw, Google may be announcing Circles today.



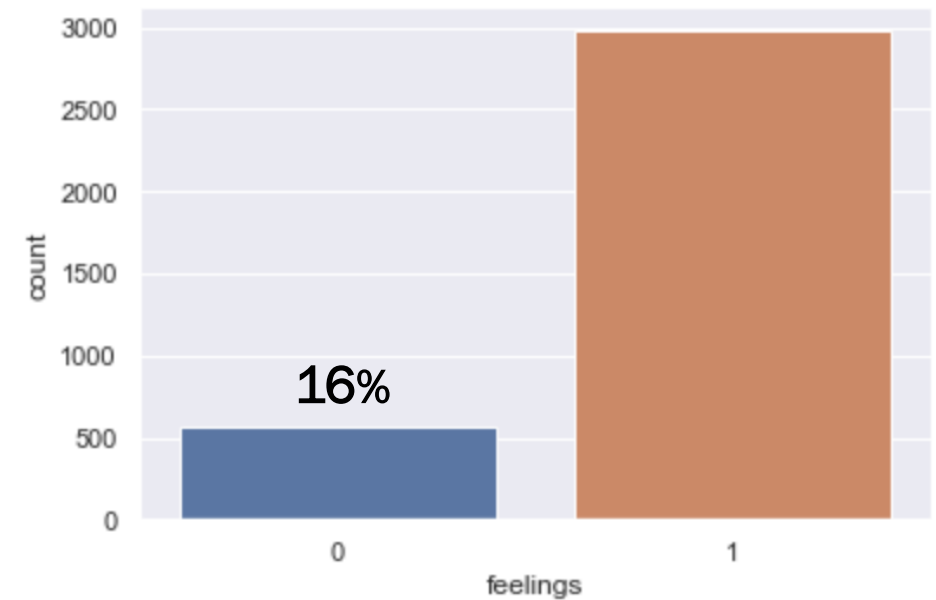
140

Tweet

9,093 Tweets

Human classified:

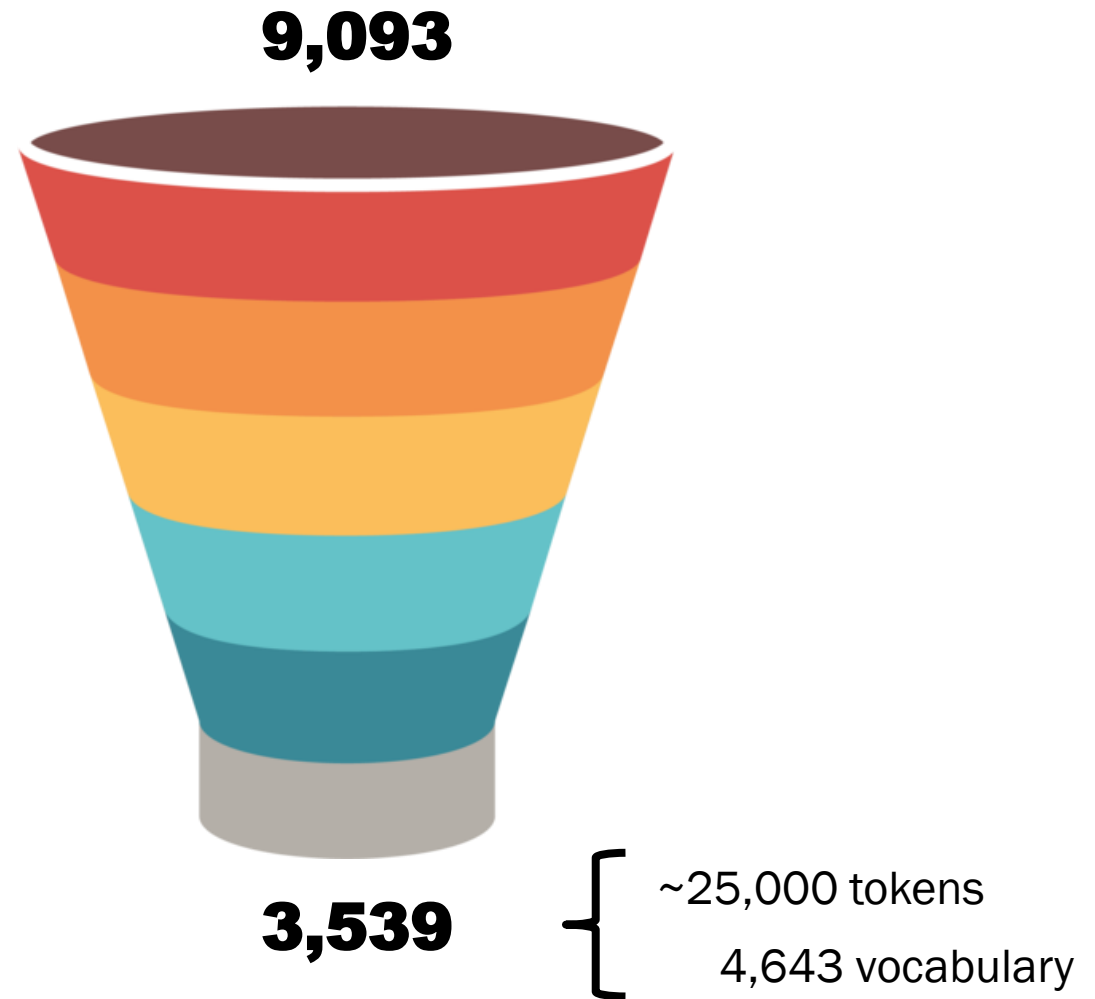
- Sentiment
 - Positive (2,978)
 - Negative (570)
 - None (5,389)
- Directed at Apple or Google
 - Specific product / app (2,200)
 - Company (1,091)



Data source: <https://data.world/crowdflower/brands-and-product-emotion>

Clean & Shape the Data

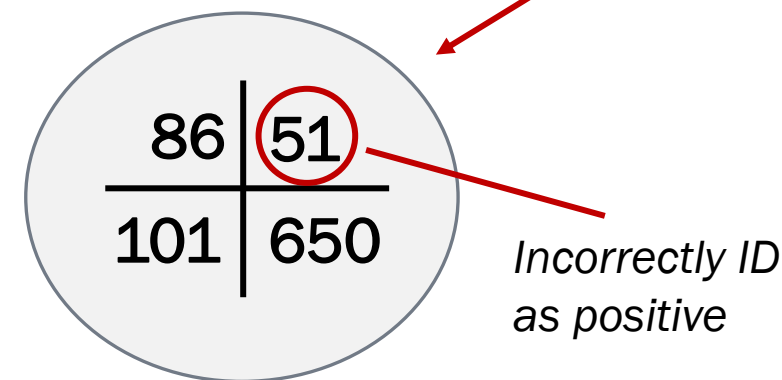
- Address missing & duplicates
- Simplification of classes
 - Positive or Negative
 - Google or Apple
- NLP processing
 - Lowercase words
 - Strip out punctuation
 - Remove stop words (211)



COUNT VECTORIZER	Accuracy	Precision	Recall	F1	TF-IDF	Accuracy	Precision	Recall	F1	LEMMAZIZED CV	Accuracy	Precision	Recall	F1	LEMMAZIZED TF-IDF	Accuracy	Precision	Recall	F1
Logistic Regression	0.84	0.51	0.42	0.46		0.87	0.63	0.54	0.58		0.86	0.54	0.42	0.47		0.87	0.57	0.51	0.54
Random Forest Vanilla	0.87	0.71	0.31	0.43		0.88	0.89	0.32	0.47		0.89	0.81	0.32	0.46		0.89	0.77	0.35	0.48
Multinomial Naive Bayes	0.86	0.65	0.28	0.4		0.85	1	0.08	0.15		0.88	0.71	0.3	0.42		0.86	0.88	0.11	0.19
SMOTE					AND SMOTED					AND SMOTED					AND SMOTED				
Logistic Regression	0.81	0.46	0.59	0.52		0.88	0.82	0.33	0.47		0.81	0.42	0.56	0.48		0.87	0.59	0.53	0.56
Random Forest Vanilla	0.77	0.39	0.65	0.49		0.88	0.86	0.23	0.37		0.79	0.38	0.55	0.45		0.9	0.84	0.4	0.54
Multinomial Naive Bayes	0.84	0.54	0.54	0.54		0.8	0.41	0.63	0.49		0.84	0.49	0.61	0.54		0.83	0.45	0.62	0.52

An iterative, layered approach...
total of 24 models run:

- Best Accuracy = **0.90**
- Best Recall = **0.65**



Best Model: a compromise of Accuracy & Recall

- Multinomial NB w/TFIDF vectorized + lemmatized + SMOTE = **0.83** accuracy
- With recall of **0.62** for class 0, there is still a 38% chance of predicting positive when tweet is negative

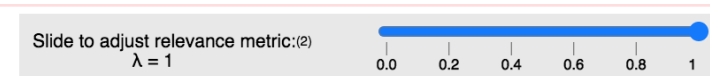
Words Contributing the Most

Important to RF models:

- Top words mostly negative
- Long tail...many words contribute a little
- *Can be difficult to glean meaning w/out context*

Word	Importance
fail	0.0179
headache	0.0148
battery	0.0145
long	0.0136
need	0.0117
think	0.0099
design	0.0094
suck	0.0092
people	0.0089
yet	0.0085

Selected Topic: 1 Previous Topic Next Topic Clear Topic



Inter-topic Distance Map (via multidimensional scaling)

PC1

PC2

1

2

3

4

Marginal topic distribution

Term	Before (%)	After (%)
circles	7.5	8.5
social	4.0	11.5
maps	3.5	3.0
title	3.5	0.5
tag	3.5	1.0
images	3.0	0.5
diller	3.0	0.5
people	3.0	1.0
major	3.0	4.0
graph	2.5	0.5
open	2.5	0.5
service	2.5	2.0
it	2.5	0.5
like	2.5	3.0
needs	2.5	3.0
better	2.5	1.5
network	2.5	6.5
launch	2.5	8.0
called	2.5	3.0
product	2.5	3.5
instead	2.5	1.5
seo	2.5	1.0
tv	2.5	3.0
get	2.5	2.0
trajan	2.5	0.5
protocol	2.5	0.5
gt	2.5	1.5
destroyed	2.5	0.5
clean	2.5	0.5
websites	2.5	0.5

Top positive words for apple (on the left) – Word cloud based on token count
Top negative themes for Google (on the right) – via LDA clustering

Out[33]:

	text	brand	feelings	text_processed
4	@sxtxstate great stuff on Fri #SXSW: Marissa M...	Google	1	@sxtxstate great stuff on fri #sxsw: marissa m...
5	#SXSW is just starting, #CTIA is around the co...	Google	1	#sxsw is just starting #ctia is around the cor...
8	Excited to meet the @samsungmobileus at #sxsw ...	Google	1	excited to meet the @samsungmobileus at #sxsw ...
9	Find & Start Impromptu Parties at #SXSW Wi...	Google	1	find & start impromptu parties at #sxsw wi...
10	Foursquare ups the game, just in time for #SXS...	Google	1	foursquare ups the game just in time for #sxsw...

In [74]: data4.text.sample(20)

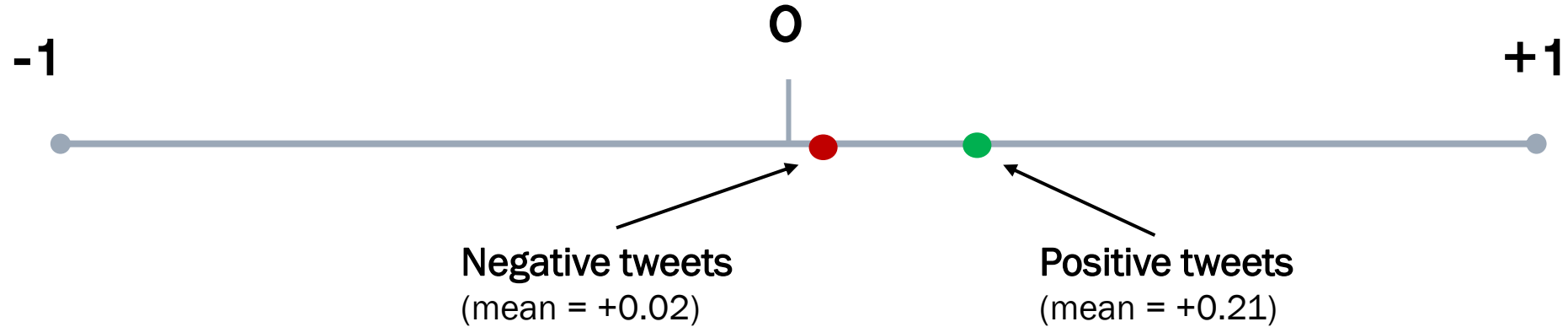
Out[74]: 1221 "Google before you tweet" is the new "think before you speak." - Mark Belinsky, #91ltweets panel at #SXSW.
2818 At #SXSW seeing a demo of #Google maps for mobile 5.2. 3D rotational viewing is very cool
1137 Mayer comes out sans intro, still gets cheers. #techrockstar Launches into Google's priority on location - Fast, Fun & Future #sxsw
2139 RT @mention Google's Marissa Mayer on the location-based 'fast, fun and future' {link} #SXSW #SXSWi
2096 RT @mention Geeking out on YouTube APIs #SXSW @mention Google Teaching Theatre {link}
925 ballroom d: #marissagoogles talking about some cool projects (obv). love the Google Art Project. #sxsw
2987 P.S. @mention and Google throw a b!tchin' party. Shout out to The Spazmatics #sxsw
250 Loved the honesty in Google's Marissa Mayer keynote: we have too many products and need to step up customer service for locations #sxsw
2130 RT @mention Google to Launch Major New Social Network Called Circles {link} #sxsw / cc @mention @mention
2489 Leaving Google's Marissa Mayer Keynote. Interesting details on user adoption of location-aware services. #sxsw
2099 RT @mention Get it while it's hot! The latest version of Whrrl is available today for Android, iPhone - and Blackberry! WHRRL FOR BLACKBE
1473 Great talk on using game mechanics to get user engagement @mention from google rocked it. #sxsw
912 When brand focuses on purpose, not object, they survive & succeed. Google: not search, useful info. Nike: not sneakers, performance.
78 Just left #sxsw tradeshow demo of @mention at the Google Theatre. Ok, I get it. I see why all the presenters here are using it.
666 Watching a promo for Google earth engine at 'Techies can save the world, why don't they?'. Harnessing collective power for good. #sxsw
19 Okay, this is really it: yay new @Foursquare for #Android app!!!!!! kthxbai. #sxsw
2976 @mention this time next week Google party at #SXSW!
1084 If you aren't at google you just missed the dance party of a lifetime #SXSW
141 ...

Read texts to understand context / meaning

Seeing specific words in context... the full tweet... adds meaning to themes and word lists

AKA... know thy data!

TextBlob polarity scale



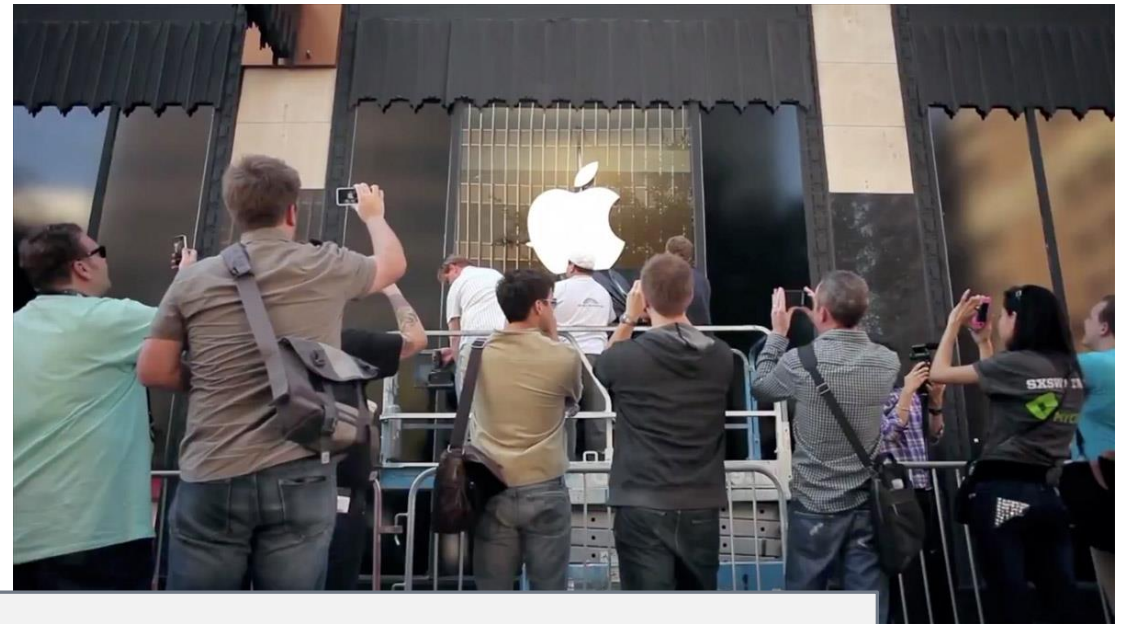
Sentiment may not be as polarized as assumed

Used TextBlob (model) to do assess how it rates polarity of sentiment in our tweets.

On a scale of -1 to +1 our negative tweets average = +0.02, and positive tweets average = +0.21

Key Findings

- Accuracy good, Recall low
 - Class imbalance
 - Tweets varied; repetitive
- Able to id key topics (+ / -)
 - Most tweets fairly neutral
 - Same word/topic can be both + and -
- Created interactive tool for applying models to data



One of the top words = **popup** or **pop...**
is it **positive** (can't wait to get my iPad2)
or **negative** (long lines) ?



Recommendations

GOOGLE

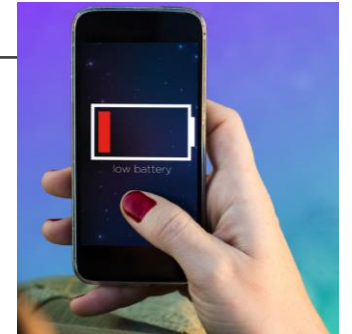


- Google Circles: too much competition
- Google Maps: fix navigation errors quicker
- Google Search: at risk from Microsoft Bing
- Company: “focus on things that matter”

Process Improvements

- **Dashboard Approach:** model tuned for speed and frequent reporting
- **Tweets:** may not be the best source of actionable feedback (part of a mix)

APPLE

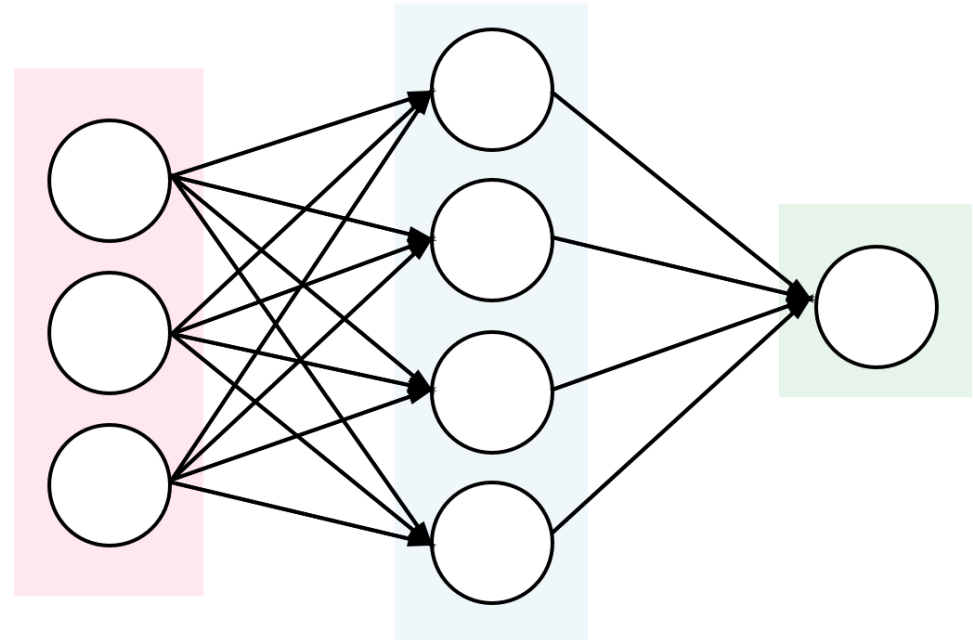


- iPhone: battery lacks longevity
- News: needs improvement and innovation
- iPad2 Launch: long lines and poor CS
- Conference: more phone charging stations



Next Steps

- **Improve Model Performance**
 - Utilize bi-grams and tri-grams
 - Filter vocabulary by parts of speech (POS)
 - Explore word embedding
 - Try using a Neural Network
- **Experiment with Content**
 - Extend data used to “neutrals”
 - Restrict tweets to specific roles / users
 - Explore different types of conferences



Thanks to...

Yish Lim

(leader of the pack)

Stephen Enke

(resident NLP expert)

Our classmates

(...and happy hour companions)

and Flatiron School



For details visit: <https://github.com/tcmcaleer/Google-Apple-Tweets>

tcmcaleer@google.com / markpatterson8@hotmail.com