



Analyzing Twitter



Twitter is a powerful tool
for businesses to engage
with their customers, but it's
impossible to manually sift
through thousands of tweets



What if we had a way to prioritize tweets for businesses to more easily engage with their customers?



1.

THE METHOD

How can we find the ratings we need for
reference?

The Method

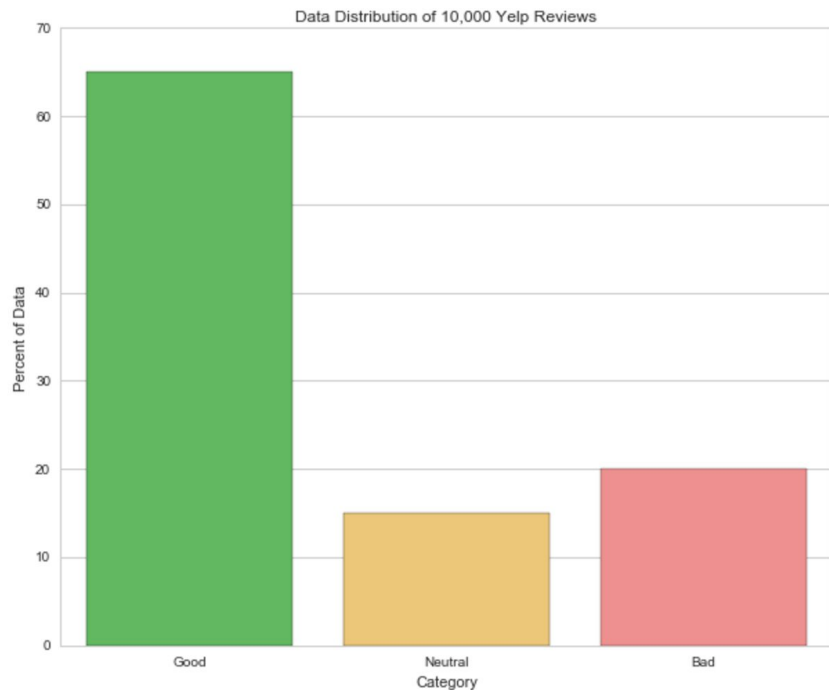


Building a
corpus with
Yelp reviews

Training various
machine
learning models

Applying optimal
model to tweets
for categorization

Data Munging



Processed 10,000 Yelp Reviews

- Spell check
- Tokenizing
- Lemmatizing

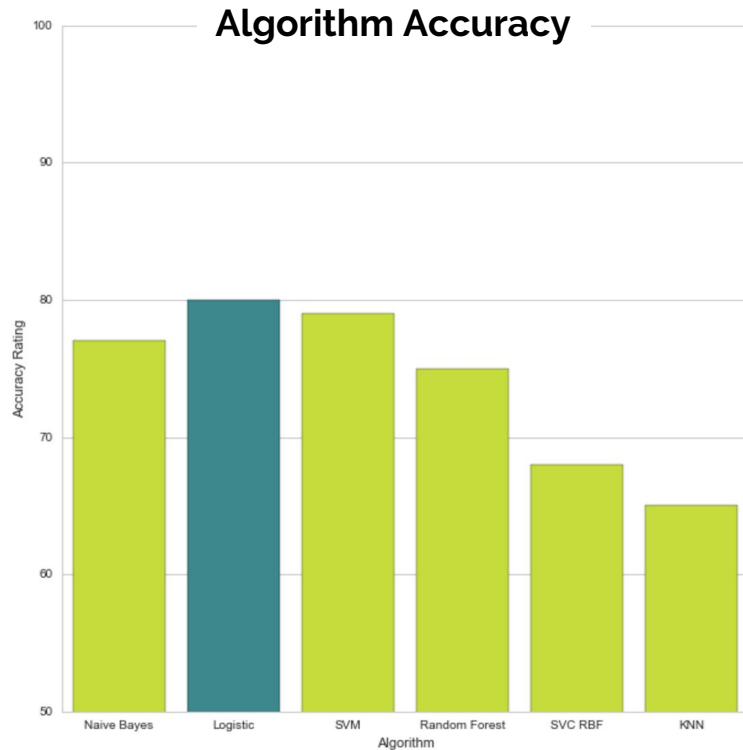


2.

THE MODEL

Applying machine learning to this problem

Modeling



Feature Generation

- Bag of Words
 - uni/bi/tri - grams
 - Simple count & Tfidf
 - Limited to 3000 features
- Sentiment Score
 - TextBlob (built on top of NLTK)

Results

Logistic Regression with TFIDF - highest accuracy 80%

	Precision	Recall	F1-Score	Test sample
GOOD	81	97	89	1646
NEUT	59	12	21	376
BAD	76	72	74	478

Model Applied to Tweets

Place	Tweet	Pred.	Prob. % four/five	Prob. % three	Prob. % one/two
Cheese Fact.	WHY YOU GOTTA FIGHT WITH ME AT [CHEESECAKE] YOU KNOW I LOVE TO GO THERE	GOOD	94%	1%	5%
Hooters	such truly beautiful ladies!	GOOD	92%	3%	4%
Chilis	so disappointed in you guys! ? you've never messed up so bad! Not only is is SO wrong, it's gross!	BAD	0%	81%	18%
Chilis	But okay thanks	NEUT	36%	40%	25%

Note: May not sum to 100% due to rounding

Cross Validation: Homo Sapiens v HAL

		Homo Sapien (Actual)			
		Total: 35	one/two	three	four/five
HAL (Prediction)	one/two				
	three				
	four/five				
		10	9	16	

Cross Validation: Homo Sapiens v HAL

		Homo Sapien (Actual)			
		Total: 35	one/two	three	four/five
HAL (Prediction)	one/two	5			7
	three		1		4
	four/five			13	24
		10	9	16	19

ACCURACY = $19/35 \sim 54\%$

Baseline would be $\sim 33\%$

Cross Validation: Homo Sapiens v HAL

		Homo Sapiens (Actual)			
		one/two	three	four/five	
HAL (Prediction)	Total: 35	one/two	three	four/five	
	one/two	5	0	2	7
	three	2	1	1	4
	four/five	3	8	13	24
		10	9	16	19

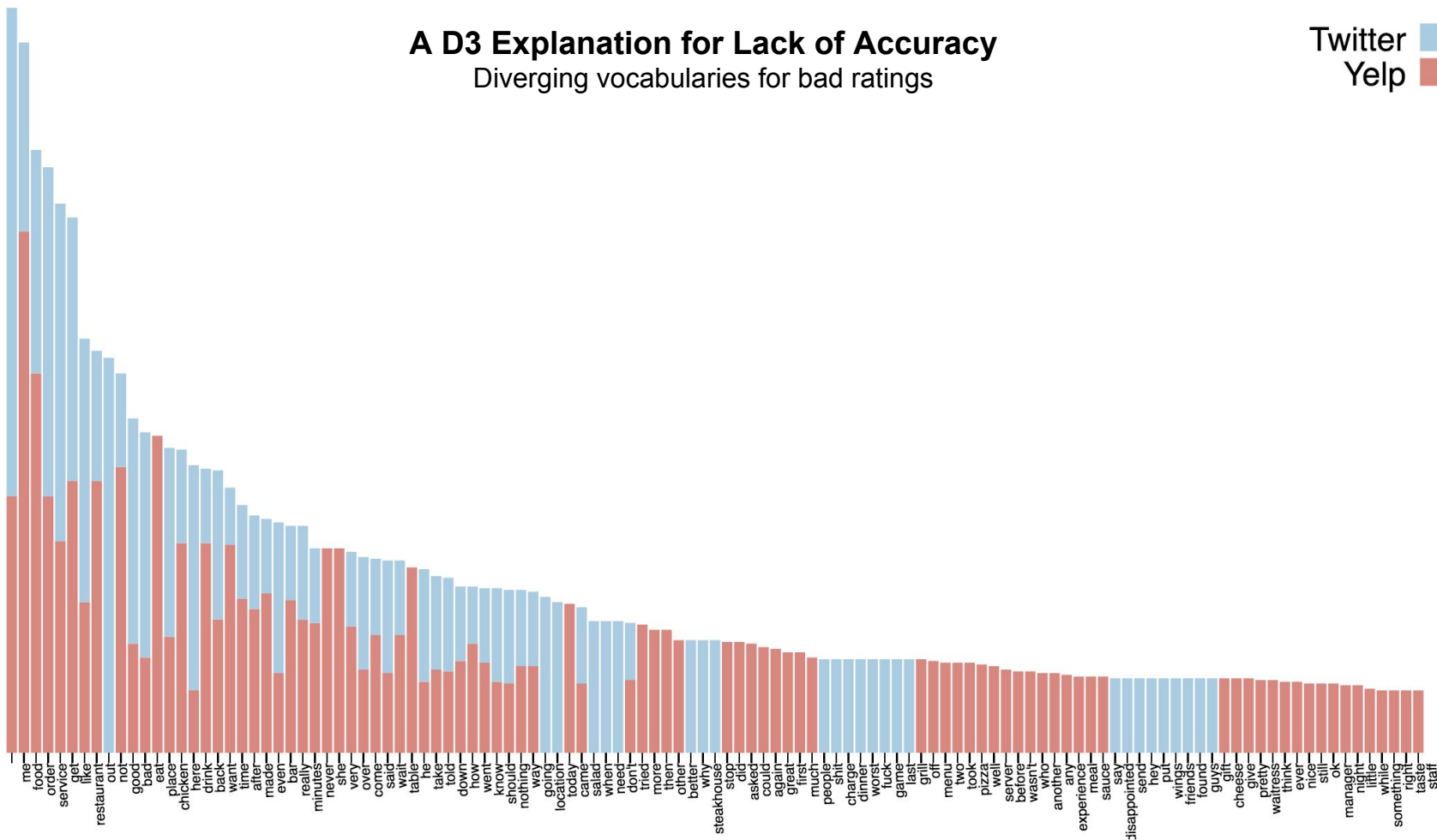
ACCURACY = 19/35 ~ 54%

Baseline would be ~ 33%

A D3 Explanation for Lack of Accuracy

Diverging vocabularies for bad ratings

Twitter ■
Yelp ■



The background consists of several overlapping geometric shapes, primarily triangles and polygons, in various shades of green and teal. The top and bottom areas are a bright lime green, while the central area is a darker teal. The shapes create a layered, mountain-like effect.

3.

THE PRODUCT

How does our final result measure up?

Demo

GO TO APP

Future Work & Other Applications

Further Development

- Increase dataset - other sites, more Yelp data
- Better processing - lemmatization, stemming, and spell checking
- Algorithm development - (e.g. neural networks)
- Advanced feature generation - (e.g. speech tagging)

Other Use Cases

- Apply to other social media channels - facebook comments, cs emails
- Raise flags on negative reviews - anomaly detection - pos/neg rate over time



Thanks!

Any questions?