

# Word Salad

Relating Food Prices and Descriptions

Victor Chahuneau

Kevin Gimpel

Bryan R. Routledge

Lily Scherlis

Noah A. Smith



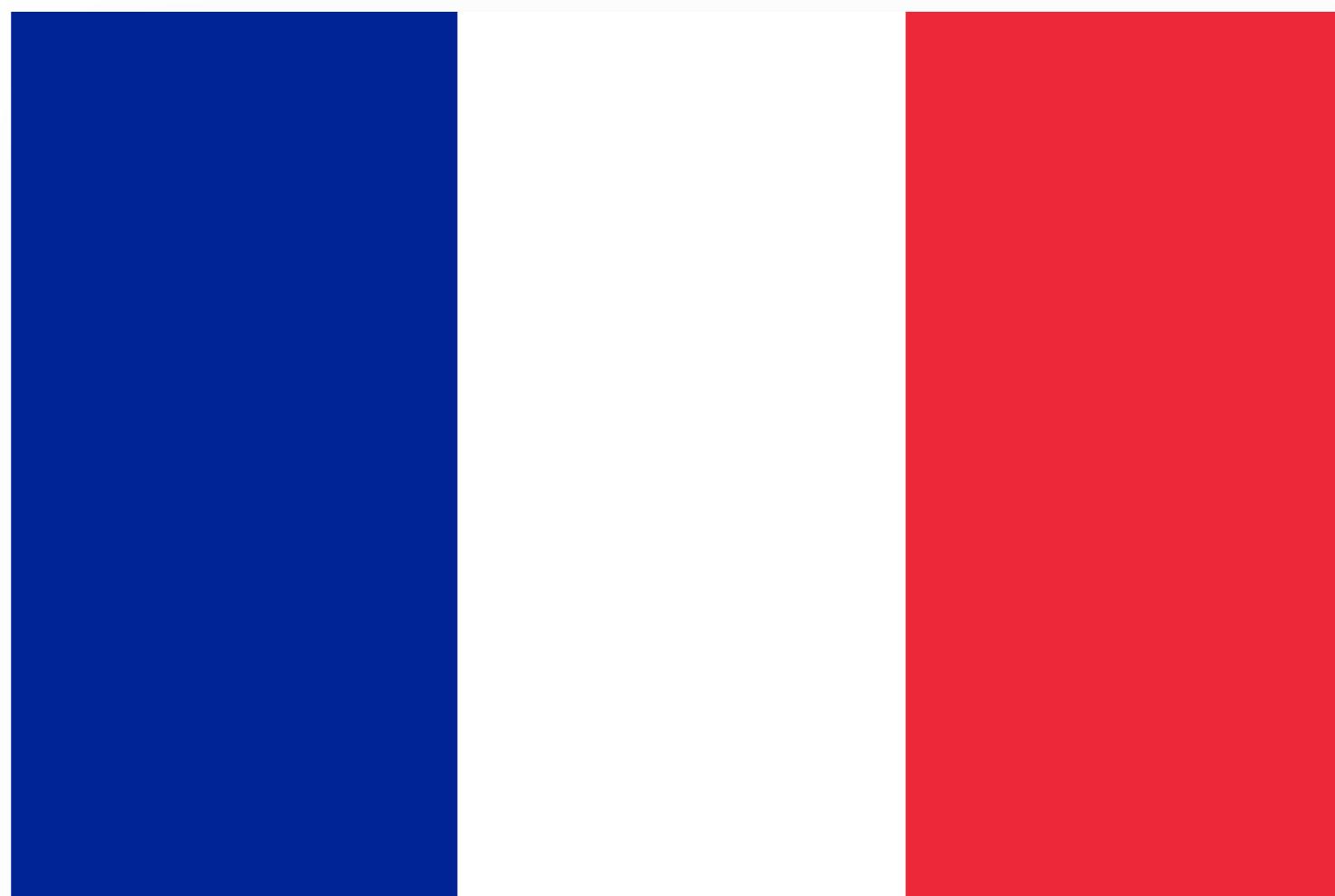
<http://bit.ly/emnlp12>

# Outline

- Motivation
  - Data
  - Models
- 1- Item price prediction
  - 2- Price range prediction
  - 3- Joint price & sentiment prediction

<http://bit.ly/emnlp12>

14 juillet 2012



# Chicken



**2 pcs fried chicken meal**

\$4.99

**Chicken Quesadillas**

Made with fresh salsa, jack and cheddar cheese.

\$6.99

**Caesar Salad**

Romain hearts, croutons, shaved parmesan cheese and classic Caesar dressing

\$9.95

**Poulet Cajun**

\$28.00

# Data



## Noah's Ark Original Deli

### Children's Menu

**MOUNT ARARAT** \$8.95  
Spaghetti & meatballs

**BIG BIRD** \$9.95  
Southern fried chicken served with french fries

**TENDER LOVE** \$10.95  
Breaded chicken tenders served with french fries

**TEENY WEENY** \$9.95  
½ Deli sandwich served with french fries

**GAN EDEN** \$10.95  
Grilled chicken breast served with french fries

### Desserts

**HOMEMADE CHOCOLATE CHIP COOKIES** \$3.95

**JELLO** \$3.95

**BROWNIE** \$4.95



## Noah's Ark Original Deli



18 reviews

Categories: [Delis](#), [Kosher](#)

399 Grand St  
(between Suffolk St & Clinton St)  
New York, NY 10002  
Neighborhood: Lower East Side

Price Range: **\$\$**

Good for Kids: Yes

Takes Reservations: Yes

Delivery: Yes

Take-out: Yes

Waiter Service: Yes

Outdoor Seating: No



Love Noah's Ark! Great comfort food. The ambiance is very deli style. Everything I've had here has been very good - from the burgers to the matzo ball soup. As a southerner, I must say it's quite exciting for me to be able to get kosher chicken fried steak. It's delicious! Their burger menu is also quite extraordinary!



Ever have a super indecisive moment when faced with a somewhat extensive online menu and needing to order food?? I feel like it's so much pressure and I usually order the wrong thing half the time. Sometimes it's also what you're craving or feel like you're craving at that moment while faced with the daunting task...



This place is kosherific. Matisyahu is their custy and you will see a pic of him eating there on the door I don't keep kosher so I wouldn't come here for a burger or anything else that would be better with cheese. I would come here for the pastrami, corned beef, brisket, and knishes.

# Data - facts

- 7 American cities

New York

San Francisco

Chicago

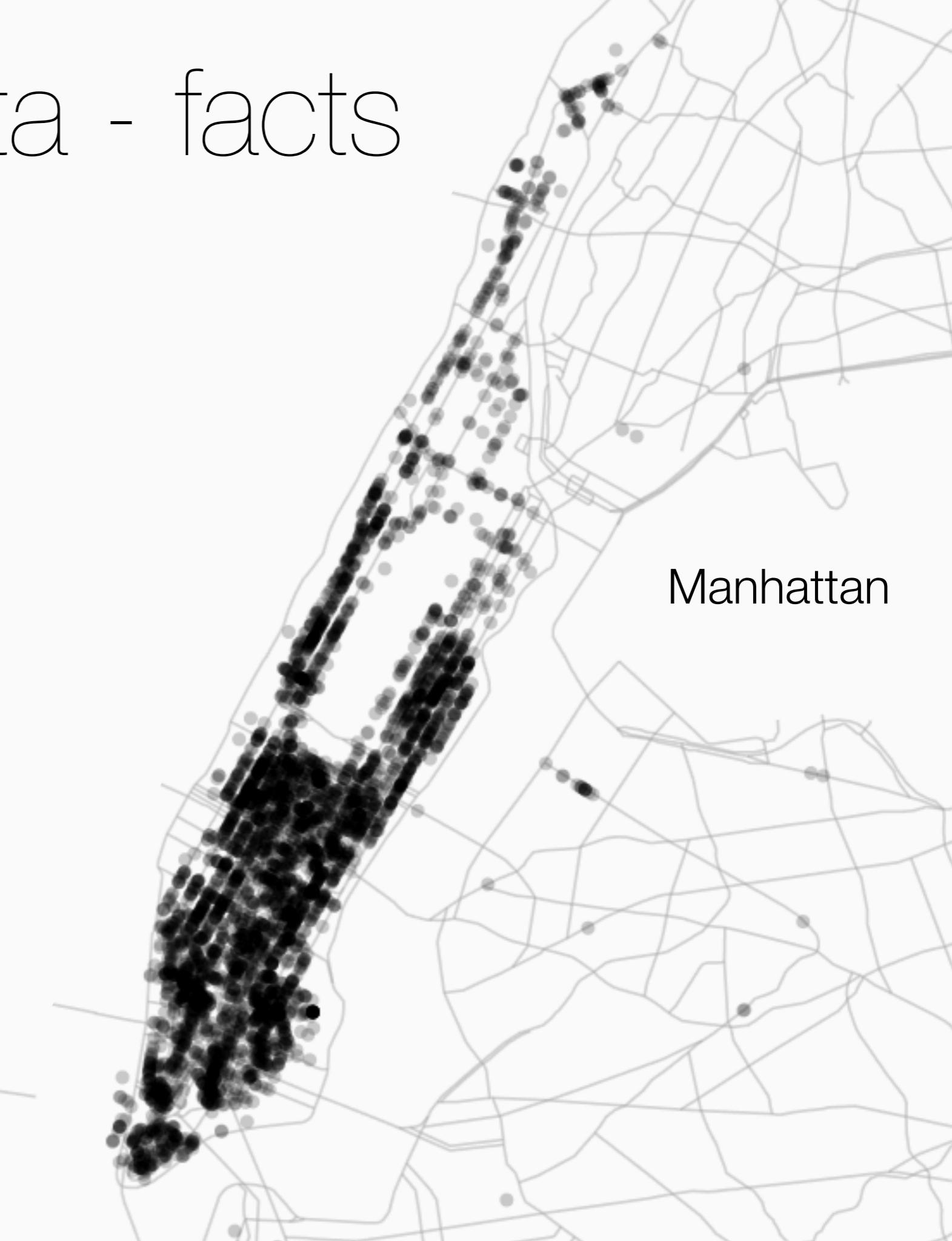
Los Angeles

Boston

Philadelphia

Washington DC

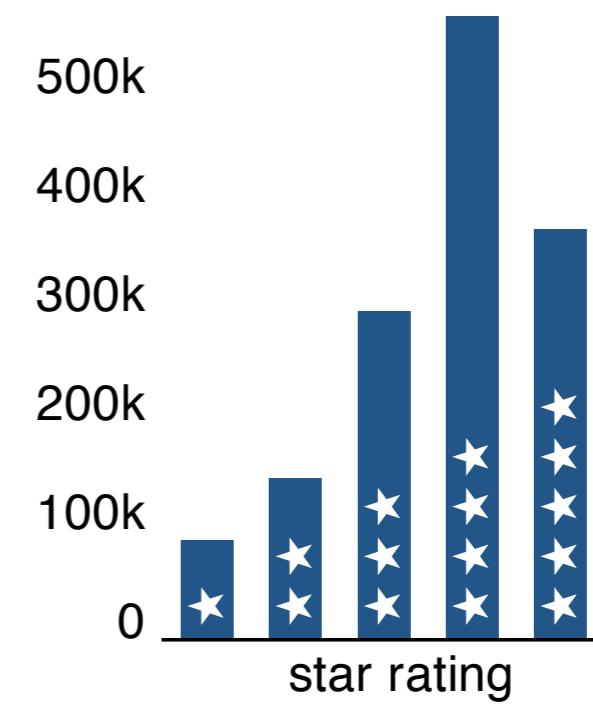
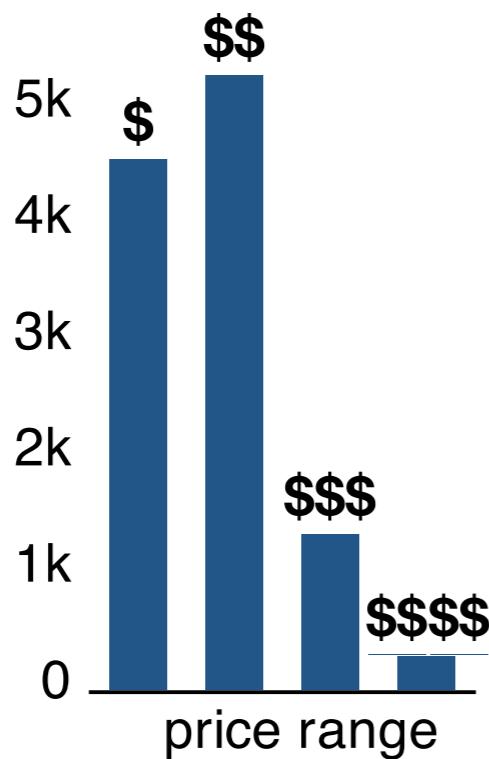
- 12k restaurants



# Data - facts

- 900k menu items  
400k distinct names

- 1.5M reviews  
80M words



# Models

- Linear
  - linear regression
  - logistic regression
- Features: binary / bag of words
- Sparsity:  $l_1$  regularization

# 1 - Item price prediction

MENU NAMES

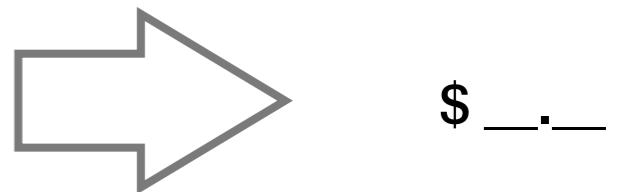
soggy fries  
n/fries+soggy  
soggy  
fries  
soggy fries

MENU DESC

drenched in brown gravy  
drenched  
brown  
brown gravy  
drenched in brown

METADATA

Category:deli  
Category:kosher  
TakesReservations:yes  
Attire:casual  
GoodForKids:yes  
HasTV:no



1-gram

2-gram

3-gram

	MAE	MRE	number features	model size
--	-----	-----	-----------------	------------

Predict mean	3.70	43.32		
Predict median	3.67	43.93		
NAME baseline	3.66	45.64	267,945	240,139
METADATA	3.55	43.11	268,450	258,828
MENU NAMES	3.23	38.33	896,631	230,840
+ MENU DESC	3.19	36.23	1,981,787	151,785
+ METADATA	3.08	34.97	1,982,363	148,774

# Authenticity in America

## Class Distinctions in Potato Chip Advertising



J. Freedman and D. Jurafsky. 2011

## MENUDESC - authenticity

		...	
old time favorite	-0.112	special	0.033
fashioned	-0.034	locally	0.037
line caught	-0.028	natural grass fed	0.038
all natural	-0.028	artisanal	0.064
traditional	-0.009	raised	0.066
natural	3e-4	heirloom	0.083
classic	0.002	wild	0.084
free range	0.004	hormone	0.085
real	0.004	farmed	0.099
fresh	0.006	hand picked	0.101
homemade	0.010	wild caught	0.116
authentic	0.012	farmhouse	0.133
organic	0.020		
specialty	0.025		
	...		

## MENUDESC - size

-0.102	slices _	}	of chicken	
-0.032	bits _			
-0.030	cubes _			
-0.024	pieces _			
-0.001	strips _			
0.015	chunks _	}		
0.025	morsels _			
0.040	pcs _			
0.042	cuts _			

## MENUDESC - style

panfried	-0.094
chargrilled	-0.029
cooked	-0.012
boiled	-0.006
fried	-0.005
steamed	0.011
charbroiled	0.015
grilled	0.022
simmered	0.025
roasted	0.034
sautéed	0.034
broiled	0.053
seared	0.066
braised	0.068
stirfried	0.071
flamebroiled	0.106

# 2- Price range prediction

## REVIEWS

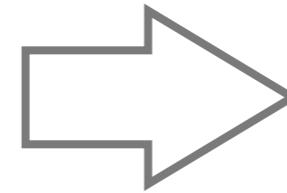
decided  
get  
cajun  
fries  
**cajun fries**  
**decided to get**  
**get the cajun**

no  
deli  
cheap  
**no deli**  
**deli is cheap**

**1-gram**  
**2-gram**  
**3-gram**

## METADATA

Category:deli  
Category:kosher  
Neighborhood:Lower East Side  
TakesReservations:yes  
Attire:casual  
GoodForKids:yes  
HasTV:no



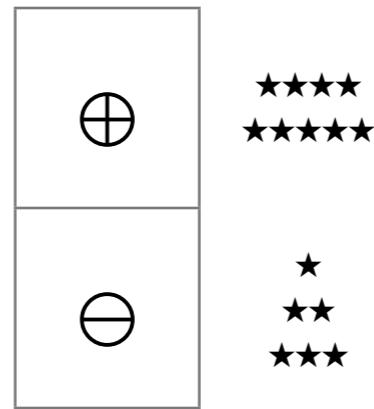
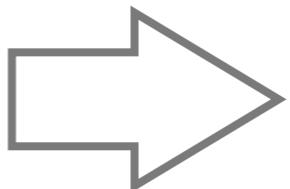
MAE	Acc.	number features	model size
-----	------	-----------------	------------

Predict mode	0.54	48.22	
MENU	0.39	66.29	1,910,622
METADATA	0.24	77.15	591
REVIEWS	0.22	79.76	3,027,470
+ METADATA	0.21	80.36	3,027,943

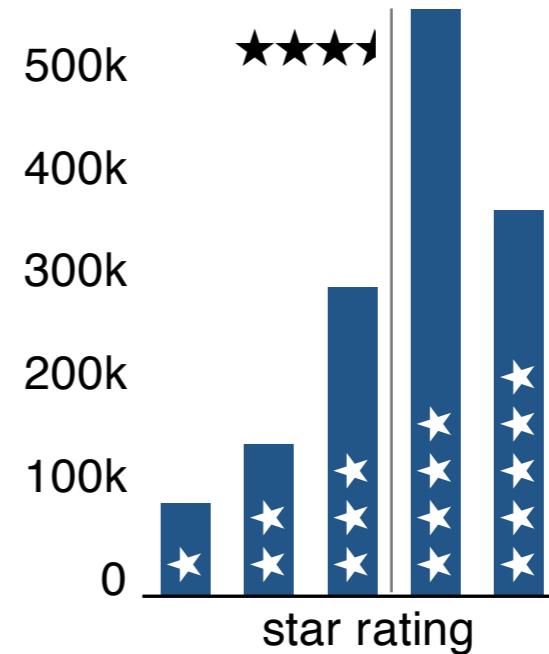
# 2- Polarity prediction

REVIEW

decided  
get  
cajun  
fries  
**cajun fries**  
**decided to get**  
**get the cajun**



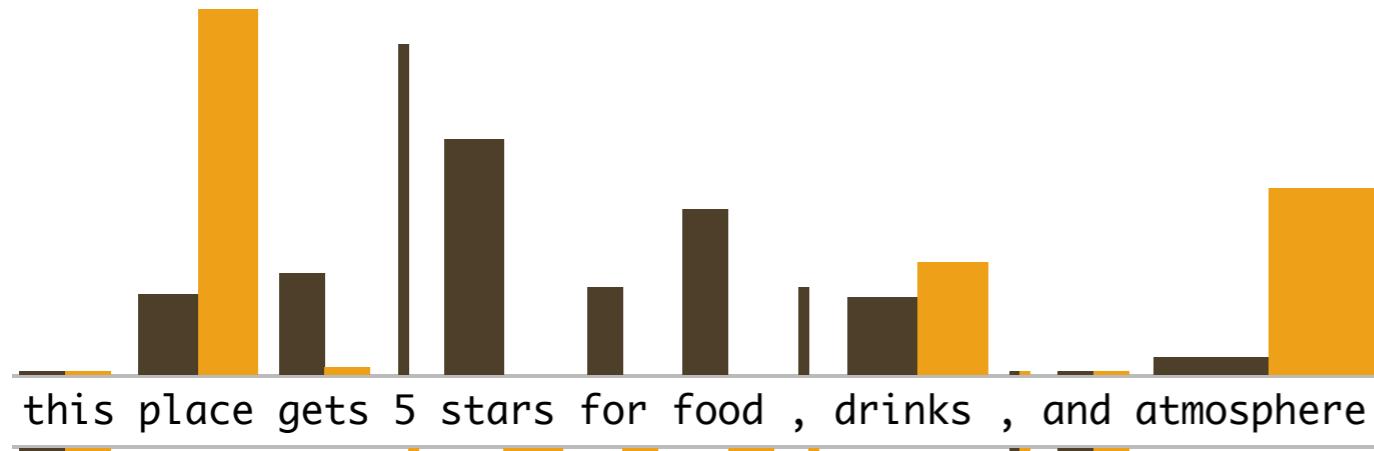
Accuracy: 87%



B. Pang and L. Lee 2008  
Opinion mining and sentiment analysis

\$ price  
★ sentiment

2/good service + 0.55  
1/service +15.13



... and negative stars for the jar-headed douche bag `` security guards ''



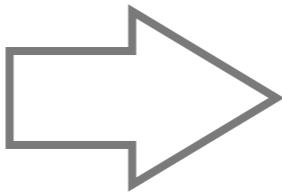
# 3- Joint \$★ prediction

## REVIEWS

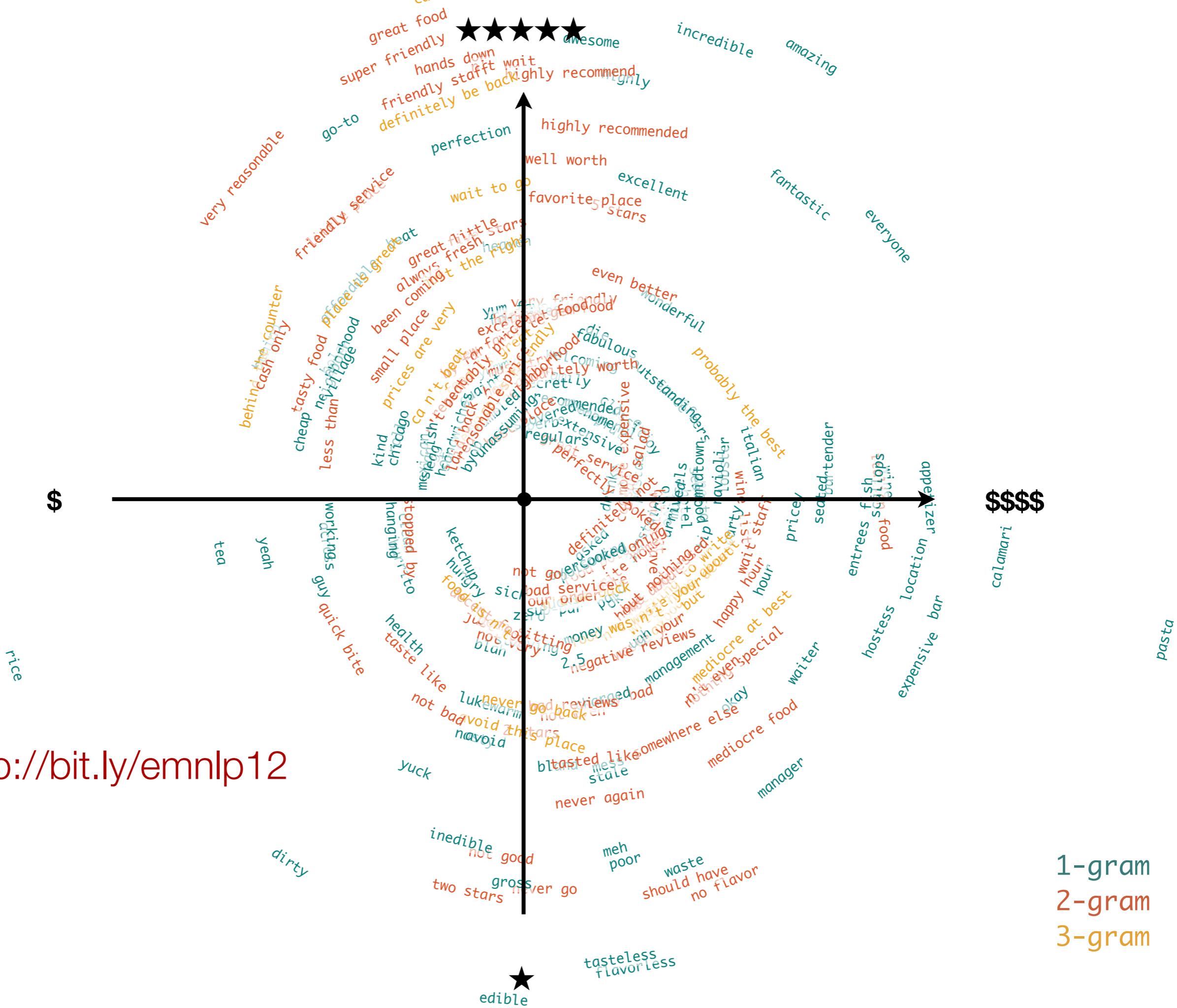
decided  
get  
cajun  
fries  
**cajun fries**  
decided to get  
get the cajun

---

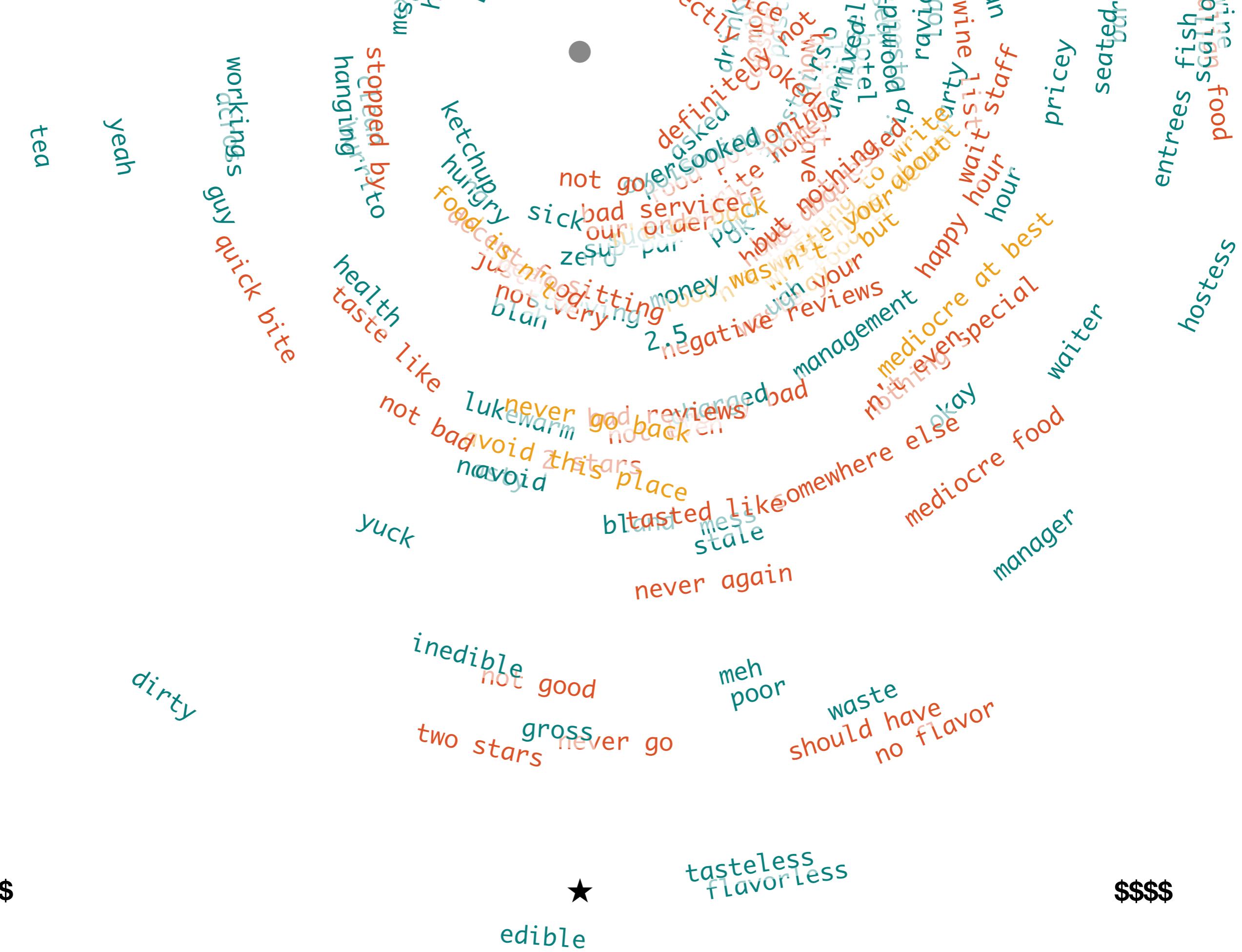
no  
deli  
cheap  
**no deli**  
**deli is cheap**



\$\$\$\$ ★	\$\$\$\$ ★★★★★
\$ ★	\$ ★★★★★







\$



\$\$\$\$

awful

disgusting

rude

edible

worse

horrible

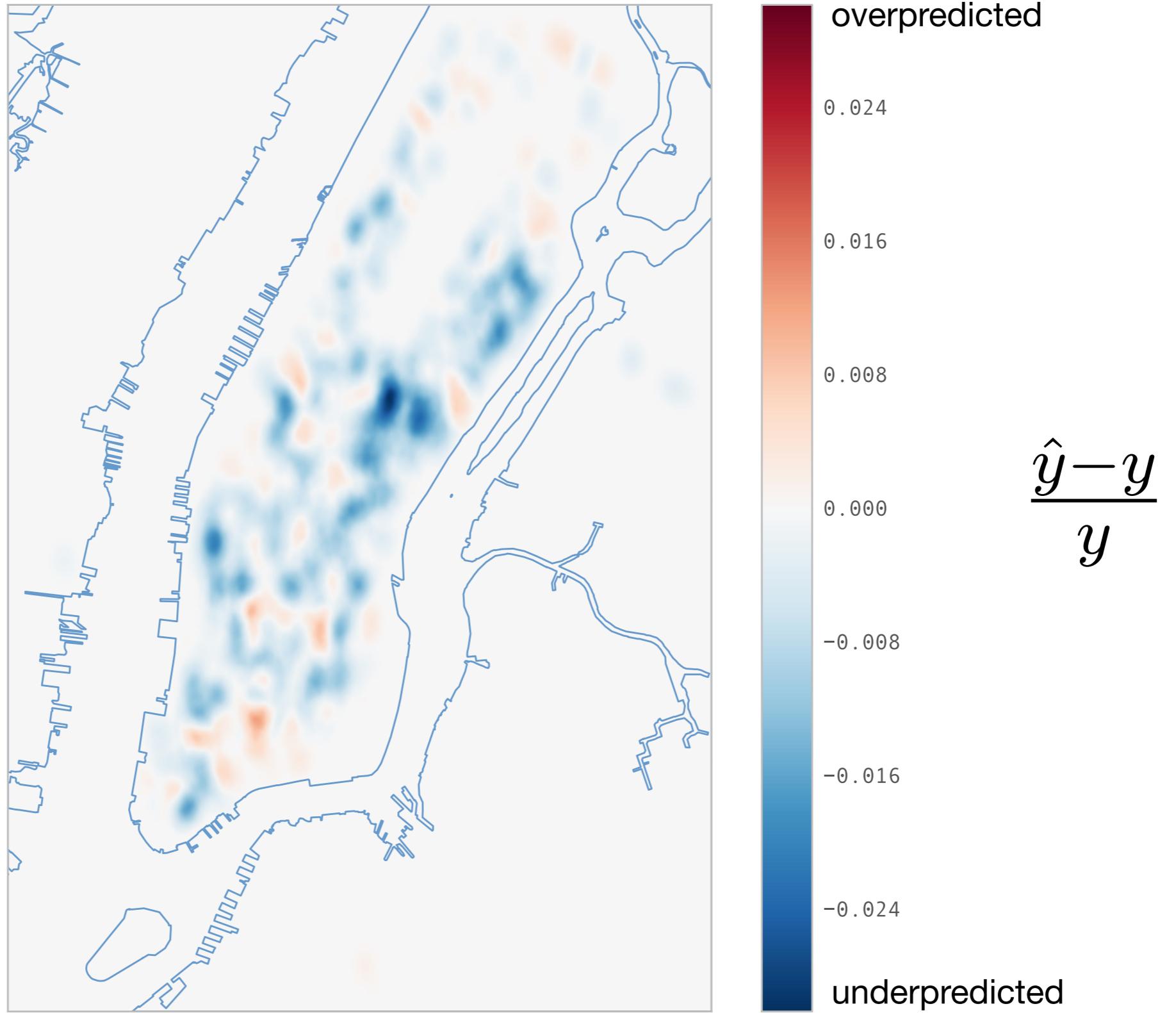
not worth

mediocre

two stars <sup>gross</sup> never go

should have  
no flavor

tasteless  
flavorless



Bon appétit !