

MINING YELP REVIEW DATA

FROM LATENT RATINGS TO A RECOMMENDATION SYSTEM

Allen Kim

YELP DATASET

Review Object

Field	Description
user_id	Unique user id
business_id	Unique business id
stars	Star rating
date	date
text	Review text

Restaurant Object

Field	Description
business_id	Unique business id
name	Business name
city	city
is_open	0 or 1 for closed or open
stars	Star rating
attributes	Business attributes
categories	Categories business belongs to

YELP REVIEWS



Jae L.

Elite '20

📷 256

✳️ 196

📝 150



9/20/2020

This place is rich with history as well as rich with delicious flavor. Ive been to Mexico a few times (from the city to southern region of Veracruz and puebla) and each place i went to, Ive always tried tacos. This place is the exact type of flavor and texture as i had back in the native land.

...

[Read more](#)



Johnny L.

Elite '20

📷 232

✳️ 472

📝 2067



9/15/2020

tried the amazing, notoriously good lamb tacos.. one question, have you had a lamb taco before? I have because I went to this place and it was tender, not too gamey.. pretty interesting

the only thing is it could've used a little more salt for the meat.. but indeed a goo...

[Read more](#)

5 STARS

4 STARS

PROBLEMS WITH THE RAW STAR RATING

1. User give rating based on number latent criterias
 - a. Opinion about aspects of restaurant experience (e.g. location, quality of food)
2. Star rating cannot reflect those criteria
3. Need a way to extract useful information about different aspects of restaurant experience from text

GOAL: Come up with a recommender system that incorporates user importance on different aspects

WORD CLOUDS OF TEXT REVIEW

1-Star Review



5-Star Review



TOPIC MODELING OF FOOD CATEGORIES



TOPIC MODELING OF LATENT ASPECTS

PRICE

Worth
Check
Price
Reasonable
Pay
Size
Small

AMBIENCE

Atmosphere
Vibe
Spot
Clean
Location
Decor
smell

SERVICE

Time
Service
Slow
Fast
Bartender
Seat
Waitress

FOOD

Quality
Order
Appetizer
Delicious
Chicken
Dish
Meal

TOPIC MODELING OF LATENT ASPECTS



Ambience



Service



Pleasant dinning area with confidence attentive servers and beautifully presented fare.

I have always found the Toro & Maguro very fresh here. The smoked Salmon is always delight as well as the oysters.

Food



The Tempura is plentiful yet never greasy.

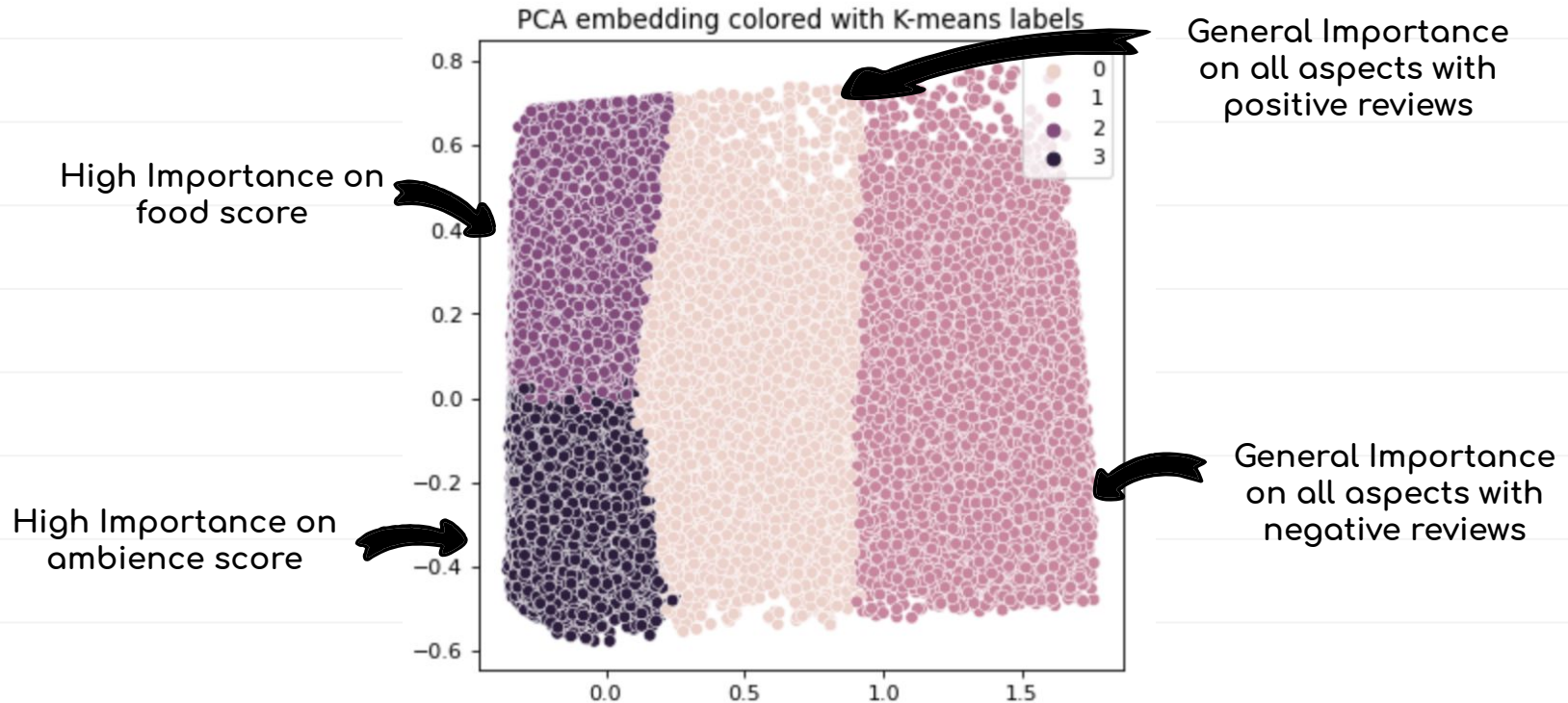
Food: 0.52

Service: 0.25

Ambience: 0.22

Price: 0

K-MEANS CLUSTERING AND PCA



RECOMMENDATION SYSTEM

1. User picks a type of food from 15 categories
2. User ranks the 4 criteria (aspects) on a scale of 0-5 for importance
3. Recommend Top N restaurants

LOOKING AT SPECIFIC USER

Rudy's Sub

Sandwiches,
American

"The food here is so great that I won't eat subs anywhere else. It's family owned and the staff is awesome. If you haven't tried this place yet, you're definitely missing out."

Food: 5

Service: 5

Ambience: 0

Price: 0

RECOMMENDED TOP 5 AMERICAN RESTAURANTS

BISTRO 19	BISTROS, AMERICAN	FOOD: 0.52	SERVICE: 0.4	AMBIENCE: 0.36	PRICE: 0.14
DIVE BAR AND GRILLE	BURGERS, AMERICAN	FOOD: 0.44	SERVICE: 0.39	AMBIENCE: 0.47	PRICE: 0.19
PATRICK'S PUB AND GRILLE	AMERICAN, BARS	FOOD: 0.57	SERVICE: 0.53	AMBIENCE: 0.25	PRICE: 0.2
SAVOY	AMERICAN, BARS	FOOD: 0.49	SERVICE: 0.41	AMBIENCE: 0.39	PRICE: 0.14
SUNSERI JIMMY + NINO	AMERICAN, SANDWICH	FOOD: 0.51	SERVICE: 0.51	AMBIENCE: 0.38	PRICE: 0.15

THANK YOU!

