

Building a Yelp Recommender System

Tobias Starling, Deanna Hu,
Nabeel Vakil, Derek Gong

Group 14



Overview



Goal

Given a restaurant or user, recommend a list of ten restaurants.



Data

Collected from an [Open Source Yelp API](#)



Natural Language Processing

Used NLP to process written reviews and refine recommendations.



Implementation

Built a collaborative filtering recommendation system using KNN with cosine similarity and content based filtering.

The Model

Filters:

- `str.contains("Restaurants", na=False)`
- `review_count >= 100`
- `stars >= 3`
- `is_open == 1`
- `State == 'CA'`

Variables of Interest:

- Price
- Location
- Type of restaurant
- Number of stars
- Weighted sentiment
 - Service
 - Atmosphere

User-Item matrix:

name	# 20	# 34	# 41
1114 Sports Bar & Games	0.0	0.0	0.0
AH Juice Organics	0.0	0.0	0.0
ASIE	0.0	0.0	0.0
Alcazar Tapas Bar	0.0	0.0	0.0
Alessia Patisserie & Cafe	0.0	0.0	0.0
Altamiranos Restaurant	0.0	0.0	0.0
Andersen's Danish Bakery & Restaur.	0.0	0.0	0.0
Angel Oak	0.0	0.0	0.0
Anna's Marketplace Bakery	0.0	2.0	0.0
Apna Indian Kitchen	0.0	0.0	0.0

Goal: Give a list of ten restaurant recommendations based on either the specific user or a given restaurant

KNN with cosine similarity:

```
NearestNeighbors
NearestNeighbors(algorithm='brute', metric='cosine')
```

Collaborative Recommender Based on a User:

- Created a user-item matrix
- Used a CSR matrix to handle sparsity
- Fit the matrix into a KNN model with cosine similarity as the metric

Content-Based Recommender Based on a Given Restaurant:

- Used NLP and other numeric features to calculate cosine similarity
- Assessed quality of service and atmosphere to create the weighted sentiment

Results

Recommendations for user 23320:

No		Name	Distance	Average Rating
1	1	Finch & Fork	0.517765	4.311321
3	2	Brophy Bros - Santa Barbara	0.532178	4.018018
23	3	Brophy Bros - Santa Barbara	0.562104	4.018018
4	4	Santa Barbara FisHouse	0.572573	4.037500
24	5	The Shop Kitchen	0.589906	4.359551
41	6	Finch & Fork	0.622852	4.311321
16	7	Scarlett Begonia	0.630092	4.000000
18	8	Corazon Cocina	0.653657	4.505263
28	9	East Beach Tacos	0.655732	4.357143
33	10	Brasil Arts Cafe	0.663182	3.948718



Finch & Fork

Brophy Bros



Santa Barbara
Fish House

Results

Recommendations for Benchmark Eatery enjoyers:

No	Restaurant Name	Similarity Score	Stars
1	Hollister Brewing Company	0.999996	3.500000
2	Shoreline Beach Cafe	0.999991	4.000000
3	The Honor Bar	0.999986	4.500000
4	The Blue Owl	0.999983	4.500000
5	The Natural Cafe	0.999962	4.000000
6	State & Fig	0.999945	4.000000
7	Longboard's Grill	0.999930	3.000000
8	Garrett's Old Fashioned Restaurant	0.999894	4.000000
9	On The Alley	0.999851	4.000000
10	Norton's Pastrami & Deli	0.999736	4.500000



Recommendations for Super Cuca's Taqueria enjoyers:

No	Restaurant Name	Similarity Score	Stars
1	Reyes Market	0.999895	4.000000
2	El Taco Grande	0.999460	3.000000
3	Cristino's Bakery	0.999210	5.000000
4	Cesar's Place	0.999200	4.000000
5	Yona Redz	0.997335	4.500000
6	Rudys	0.997237	3.000000
7	Altamiranos Restaurant	0.997080	4.000000
8	Pollofino's	0.996783	4.000000
9	Santa Barbara Chicken Ranch	0.996760	3.500000
10	La Guerrerita Mexican Food	0.996449	4.000000



Key Takeaways

- Both recommender systems worked well when given either a user or restaurant
 - Pain points along the way:
 - Refining the relevance of recommendations
- Future improvements:
 - Experimenting with different models
 - Expanding the scope of our model to cover the entire US
 - Optimize this for actual use in an app or on a website
- Questions?

