

Assignment #3

Gaurav Lalwani
Sonakshi Garg
Shweta Mane
Milind Siddhanti
Sanchit Singhal
Pedro Seguel

Fall 2018, November 2, 2018

MIS 381N – User Generated Content Analytics w/ Dr. Anitesh Barua

McCombs School of Business, the University of Texas at Austin

i)

The team members are the following:

Gaurav Lalwani
Sonakshi Garg
Shweta Mane
Milind Siddhanti
Sanchit Singhal
Pedro Seguel

ii)

We searched for “Vegetarian food” in the location of “Austin” and collected a total of 19,425 reviews from Yelp.

It should be mentioned that other more specific search queries were considered, such as “Vegetarian-friendly” restaurants, however, these type of options didn’t provide the minimum amount of reviews required.

iii)

Using our similarity and sentimental analysis, we recommended the following three restaurants for a customer that has specified service, food, price, and location as desirable attributes:

Recommendation Order	Restaurant	Similarity	Average Sentiment	Weighted Score
1	Odd Duck	0.78277729	0.9768	0.8798
2	My Thai Mom	0.78856435	0.9523	0.8704
3	Kerbey Lane Cafe	0.84251149	0.8935	0.8680

iv)

After scraping 19,425 reviews from Yelp for “Vegetarian food” in “Austin”, we calculated the similarity scores for each review based on a customer’s desired attributes of “service”, “food”, “price”, and “location”. Sorting it from high to low, we extracted the top 200 most similar reviews. After that, we performed a sentimental analysis on the 200 reviews and using a pivot table in Excel, calculated the average sentiment for each of the 61 unique restaurants. Now that we had a similarity and average sentiment score for each restaurant, we had the option of recommending the restaurants by ordering by the sentiment alone. This would make sense since we have already determined these reviews to be similar to the desired attributes. However, the team determined that it would make more sense to create an average score

between similarity and average sentiment so that we are suggesting a balance between the two. For this reason, a weighted score was added to the table, and we recommended the top three by sorting this column. The restaurants above resulted in the highest weighted score. When simply sorting by the average sentiment score, two of the three restaurants remain in the top three - Odd Duck and My Thai Mom. Kerbey Lane Cafe falls out and is replaced by Barley Swine. Although both approaches result in similar results - we believe by weighing both similarity and average sentiment gives us a more robust recommendation system.

We also calculated the average rating (on a scale of 5) for all 77 restaurants mentioned in the original 19425 reviews. By simply sorting highest to lowest, the top three recommendations through this approach were the following:

Recommendation Order	Restaurant	Average Rating	Count
1	Wanderlust Food Trailer	5	10
2	Chicken Lollypop	4.94	70
3	Otherside Deli	4.94	68
4	Parisian Crepes	4.93	66

We included the count of reviews for each restaurant because the top result when ordering by rating only had 10 reviews. The team believes that it is common practice to see the number of reviews for each rating as a small count can lead to misleading rating results.

When comparing the output from the two different recommendation systems, we can see an immediate difference. None of the restaurants from the two lists match up. It was interesting to note the average rating for the three restaurants we recommended through similarity and sentiment analysis:

- Odd Duck - This restaurant was ranked number 18 with 4.6
- My Thai Mom - This restaurant was ranked number 72 with 3.8
- Kerbey Lane Cafe - This restaurant was ranked number 54 with 4.0

Although the restaurants we recommended, were relatively highly rated (above 3.5), none of the three made it to the top 10 of their list.

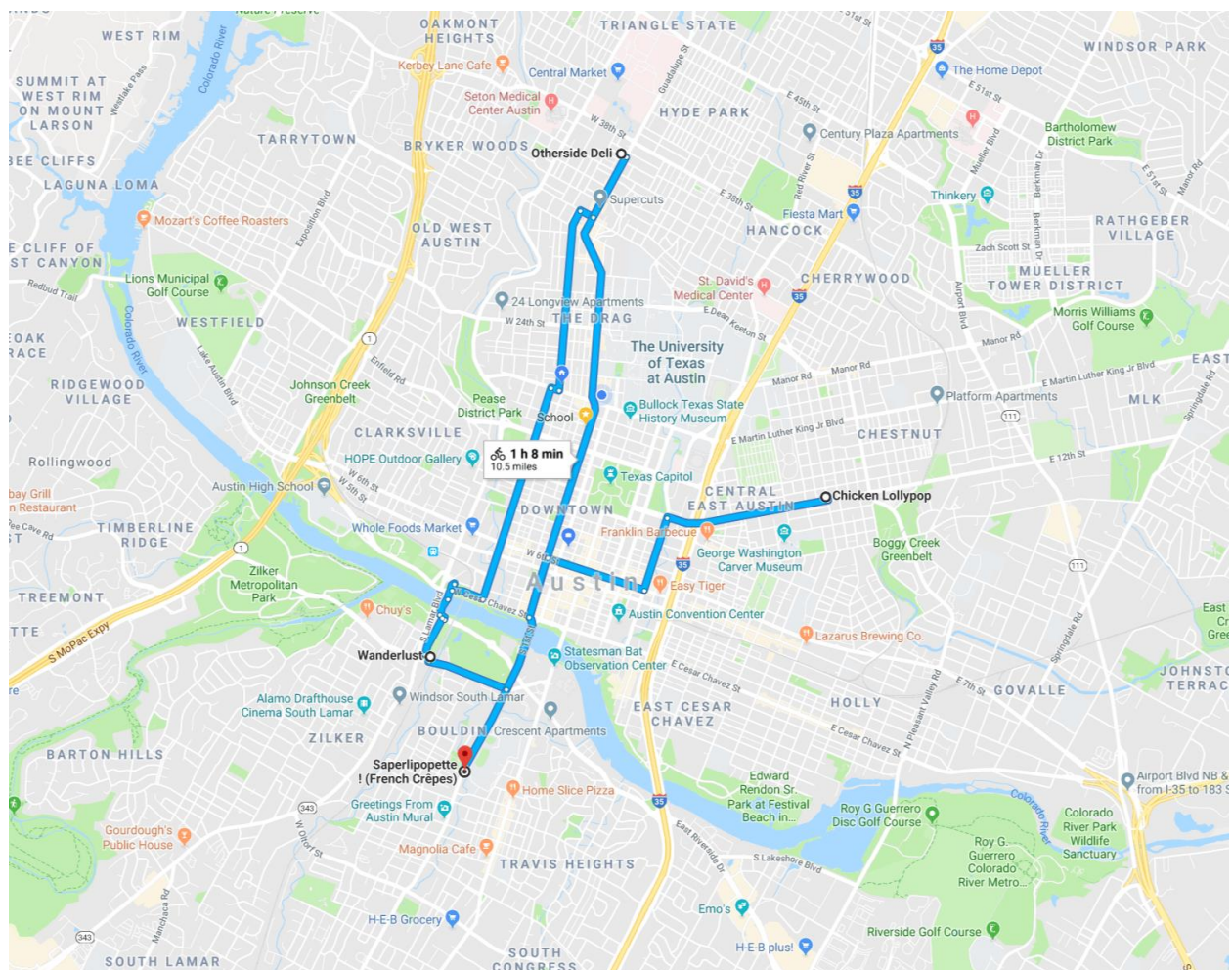
Would the top rated restaurants meet the requirements of the customer?

The user's desired attributes were service, food, price, and location. These restaurants have the highest overall rating which indicates these might have some good attributes, however, these

restaurants not coming up in the top 200 list according to similarity indicates that they lack the attributes we are searching for: service, food, price, and location.

We were a little skeptical with the second recommendation via rating (because it was called Chicken Lollypop and we were searching for vegetarian food) - but when examining their menu online, they did seem to have options for vegetarian food. However, the top-rated restaurant, Wanderlust barely seemed to have any vegetarian options - there was 1 single dish for vegetarians (eggplant parmesan) on their menu.

When investigating the location of these restaurants, it did seem like they fell within the “Austin” area and in this regard, they did well:



Checking out the menus for these restaurants, the price of the restaurants were very inconsistent. Whereas the average dish at Wanderlust costs around \$10, a meal at Chicken Lollypop would probably in excess of \$20. We were unable to determine the prices from the menu of Otherside Deli.

In conclusion, **the team does not believe the highest rated restaurant would meet the requirements of the customer.** This recommendation does a good job at meeting the location requirements (all of them were in the Austin region). However, food does not seem to match what the user desires. When looking for vegetarian options, it's unlikely to please the customer when they find out that one of the highest rated suggestions has only a single option for them. Further, the prices (one of the desired attributes), seem to be inconsistent amongst the recommendations and lacks insights that they offer to the customer. We believe our recommendation system, using similarity and sentimental analysis seeks to align with the customer's needs more accurately and would fare better for offering a personalized list of suggestions.