Topics Mentioned in Restaurant Reviews

Yelp Dataset Challenge, Round 8 (2016)

Katharina Roesler
roesler@stanford.edu

Abstract

This project examines three common topics in Yelp restaurant reviews: food, service, and money. I develop a simple method of identifying these topics and examine their relationships with review stars. I find that food is by far the most commonly discussed topic and is positively related to review stars, while service and money are negatively related to review stars.

Introduction

What makes for a great restaurant experience? Do people care only about food, or do they also evaluate a restaurant's service and monetary value? Are these topics more often mentioned in positive or negative reviews? To answer these questions, I examine 1,630,712 Yelp restaurant reviews from 26,630 restaurants. I develop measures of whether reviews mention food, service, and/or money with 98, 84, and 86 percent classification accuracies, respectively. I then use these measures to assess topics' relative frequency and importance to restaurant experiences. All code can be found here.

Supervised Topic Classification

To obtain labeled data, I labeled 165 restaurant reviews as to whether they mentioned food, service, and/or money. I find that 91 percent of reviews mention food and that 98 percent of reviews mention at least one of these topics. Below is a typical review mentioning all three topics:

> Entertaining if you have little ones - but slightly overpriced because of that. Service was attentive but but a bit slow. They appeared a little short staffed (they close at 6pm on Sundays and had a decent crowd in there right up to close).
> I had the ribs - not bad. Beer was cold; decent options for kids. -*Dan*

Next I created three regular expressions to predict whether each topic was present.

Food

To predict whether a review mentions food, I first use the following regular expression:

```
(^|\W)(food|meal|ingredients|fresh|tast|delicious|flavou?r|yum|burger|fries)
```

On its own, this measure correctly classifies 75 percent of reviews, with many false negatives overlooking particular food items. For this reason, I also use a list of 542 food and eating words from Enchantedlearning.com to predict whether reviews mention food.

Service

The next topic I consider is service, which I define to be mentioning staff, speed or waiting, or the accuracy of an order. The regular expression I use to measure service is shown below:

> (^|\W)(service|server|waiter|waitress|staff|host(ess)?\W|employee|manager|
> worker|busboy|cashier|proff?ess?ional|rude|polite|friendly|courteous|
> speed|fast|slow|time|wait|minute|hour|while|immediate|quick|
> (?<=mess up) order|(?<=screw up) order|order (right|wrong)))

Money

The last topic I examine is money, which I define as price or monetary value. I use the following regular expression to capture mentions of money:

> (^|\W)(cheap|expensive|deals?|bucks?|afford|spen(d|t)|charg|price|$|dollar|
> money|free|cost)

Results

Classification Accuracy

To validate these measures, I split the labeled data into training (70%) and test (30%) sets and predict topics using SVM, Naive Bayes, and logistic regression models. All three models have very similar test set accuracies, likely because they are very simple. The best models predict mentions of food with 98 percent accuracy, service with 84 percent accuracy, and money with 86 percent accuracy.

In general, there are more false negatives than false positives for all three topics, as the regular expressions occasionally fail to capture certain topic-related words.

Topic Distribution

Food is the most common topic and is rarely absent when other topics are mentioned. This is true for both labeled and classified reviews, as shown in Figures 1 and 2.

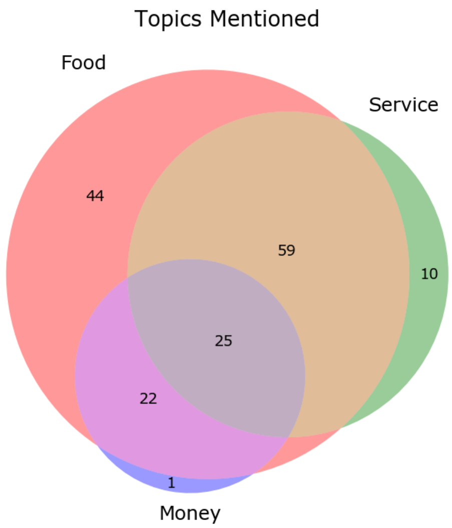

Figure 1: Labeled Reviews

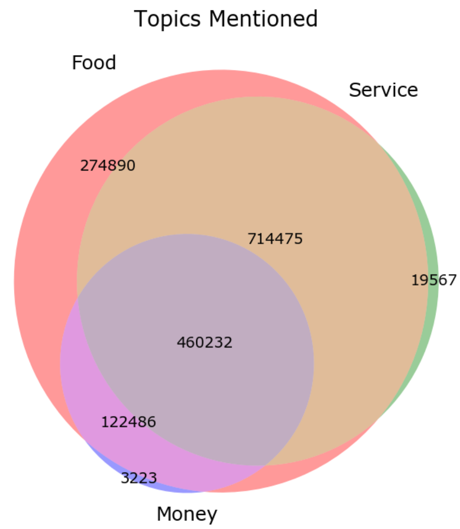

Figure 2: Classified Reviews

Topic Association with Review Stars

Positive reviews are more likely to mention food and less likely to mention service or money. Figure 3 shows the percentage of reviews mentioning each topic by number of stars, showing that one-star reviews are less likely than other reviews to mention food and more likely to mention service. In addition, five-star reviews mention money much less often than do other reviews, and the relationship between money and stars appears to be quadratic.

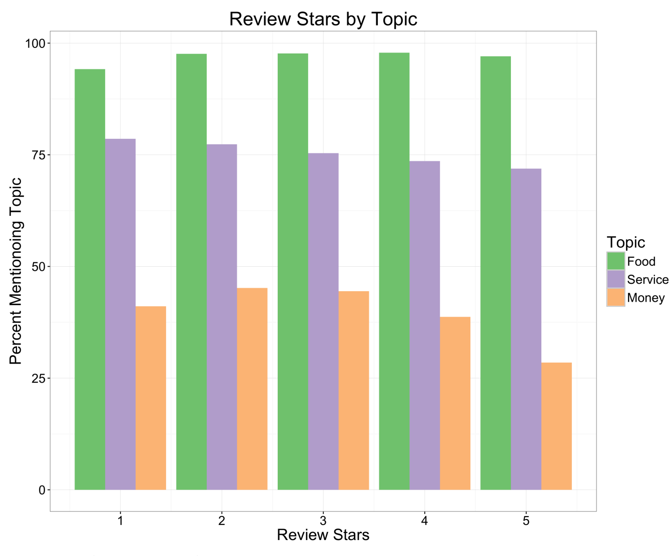

Figure 3: Topic Prevalence by Review Stars

Table 1: Linear Regression Predicting Review Stars

	Model 1	Model 2
Intercept	3.70^{***}	3.70^{***}
	(0.00)	(0.00)
Word count	-0.31^{***}	-0.29^{***}
	(0.00)	(0.00)
Word count squared	0.05^{***}	0.05^{***}
	(0.00)	(0.00)
Food		0.08^{***}
		(0.00)
Service		-0.00
		(0.00)
Money		-0.09^{***}
		(0.00)
R^2	0.029	0.037
Test Set R^2	0.029	0.036
N	1,133,026	1,133,026

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

However, it may be misleading to consider these topics in isolation from one another, as they may be correlated. For instance, mentioning money may be associated with mentioning service, and it may be money, rather than service, that affects a review's stars. In order to determine the importance of each topic net of the others, I estimate a least squares regression model predicting a review's stars (1-5) given its topics mentioned and word length. The features in these models are standardized, such that each coefficient represents the change in stars that would result from one standard deviation's increase in the feature.[1]

The results, shown in Table 1, suggest that mentioning food is positively associated with review stars, while mentioning money is negatively associated with stars. However, mentioning service is no longer associated with review stars once food and money are held constant.

Perhaps food is more often discussed in positive reviews because other topics, such as service or cleanliness, are generally taken for granted and only mentioned when something negative occurs. When this is not the case, food is the primary discussion point. And perhaps money is salient to reviewers only when food is over-priced.

Review Length

Longer reviews tend to give fewer stars, but this association diminishes as reviews get longer. This can be seen in Figure 4, which shows reviews' average stars by word count.[2]

[1]For the topic measures, this is roughly half the effect of changing from not mentioning to mentioning a topic.

[2]Nonetheless, including word count and word count squared to predict which topic is mentioned does not increase prediction accuracies.

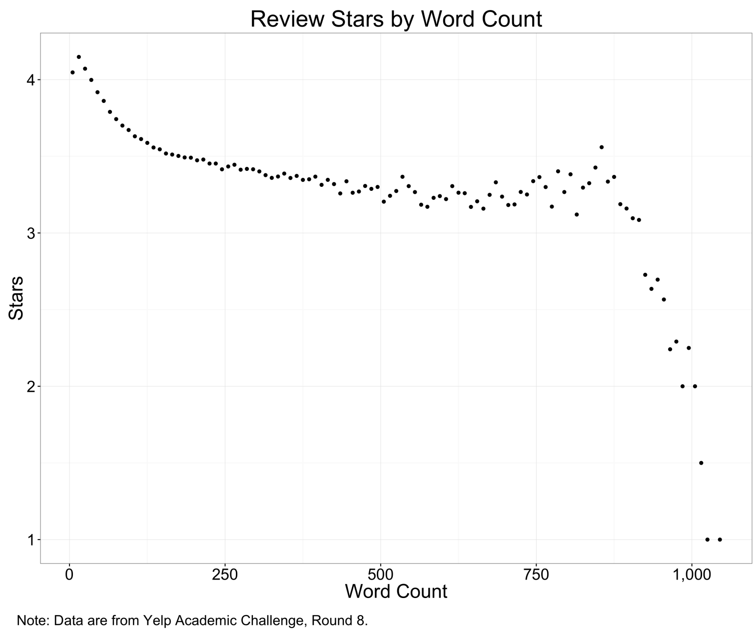

Figure 4: Review Stars by Word Count

This relationship makes intuitive sense, as people often have short positive reviews with little explanation, but tend to justify negative opinions with longer reviews. However, it is somewhat surprising that this relationship is non-linear and takes such as steep turn downwards as reviews become very long.

Unsupervised Topic Classification

Another way to examine topics mentioned in reviews is to consider words' co-occurrences. The assumption is that when words frequently co-occur with one another, they are likely to be referencing the same topic.

To examine words' relationships with each other, I create a graph with words as nodes and co-occurrences as weighted edges. For visualization purposes, I include only the 1,000 most frequent English words and the 200 most highly weighted edges.

Using a modularity-minimizing community detection algorithm provided by Thomas Aynaud, I find that three topics emerge. Figure 5 shows these word communities. The green community discusses service, while the dark blue words are about the entire restaurant experience and the light blue words are particular foods.

It is clear that food is central to restaurant experiences. For instance, the word "food" has extremely high degree centrality (0.94) and betweenness centrality (0.59). That is, 94 percent of words are in reviews also including the word "food," and 59 percent of all shortest paths between all pairs of words pass through the word "food."

In addition to this network analysis, I use latent Dirichlet allocation (LDA) to identify 20 topics and the words associated with them. However, these topics are not very meaningful and are unhelpful towards understanding the importance of food, service, and money to restaurant experiences.

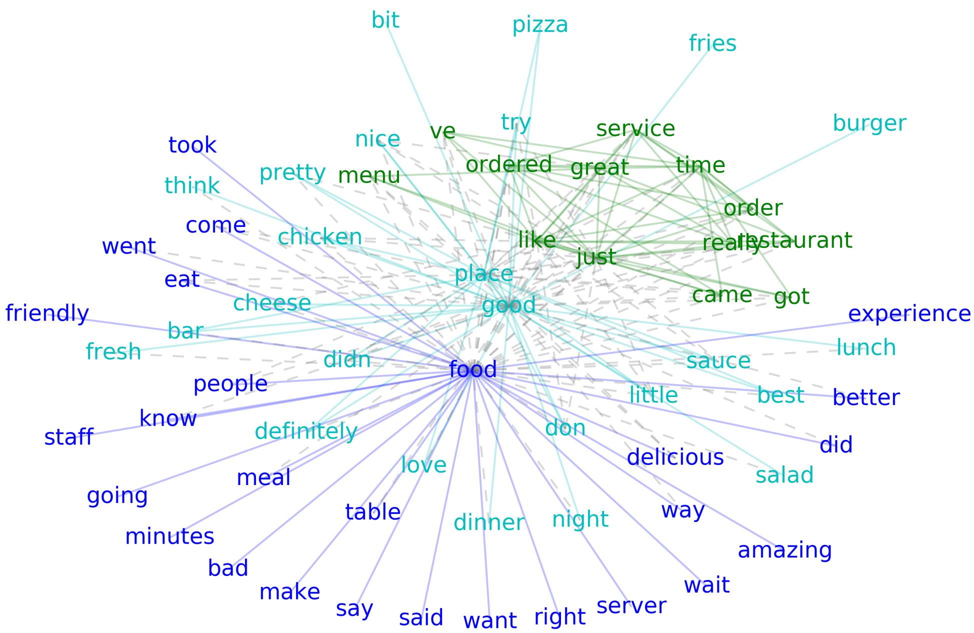

Figure 5: Word Co-Occurrences

Conclusion

This project examines three topics common to Yelp restaurant reviews: food, service, and money. It demonstrates the utility of three simple measures of these topics, as well as the importance of these topics to users' restaurant experiences.

I find that food is by far the most commonly discussed topic and is usually mentioned even when service and money are mentioned. Food is central to restaurant reviews.

In fact, mentioning food is related to having a more positive restaurant experience. On the other hand, mentioning money is related to having a more negative experience. Future research might examine why these relationships are present and classify positive and negative language within each topic. In addition, Yelp may wish to create separate scores for each topic, enabling users to assess restaurants' food, service, and value for money separately.