

# Determining Attributes of Good Restaurants Using Review Text

Analysis for  
restauranteurs/diners

Noah-John Hizon  
11/1/22

# Outline

- Business Problem
- Data Understanding
- Classifier Model
- Recommendation System
- Conclusion
- Next Steps

# Introduction

Has a restaurant ever not lived up to your expectations?

- Waiter service
- Food quality
- Atmosphere



# Business Problem

Owners with failing restaurants may be at a loss when trying to look for avenues of improvement

- New hardware, staffing, or expensive ingredients may not be worth the investment
- Difficult to isolate relevance/actionable insights from noise on a business' review page



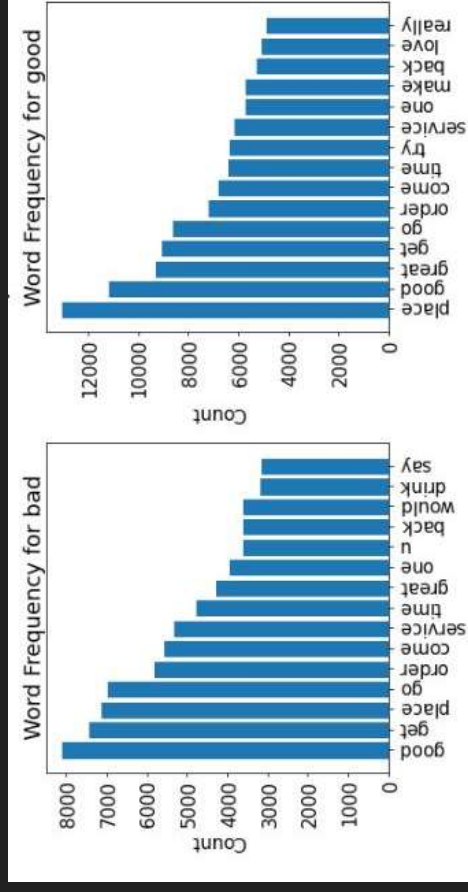
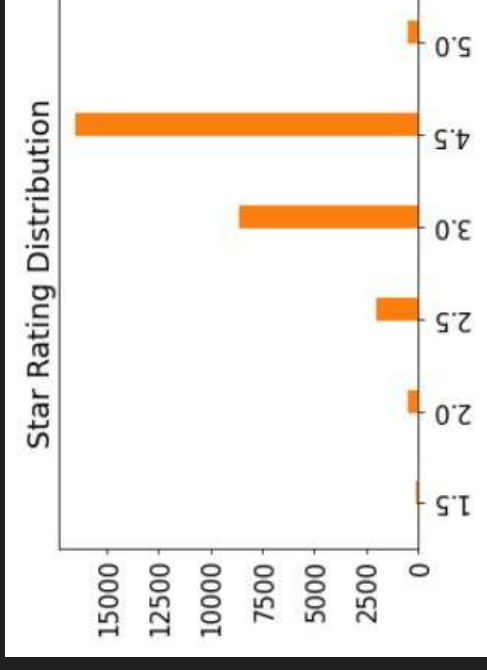
# Data Understanding

**SOURCE:** Yelp restaurant review data spanning 2010-2014 (Kaggle)

- Each row in the table represents a unique review written by a user for a specific business, with text stored in a singular string
- **Columns of interest: ‘text’, ‘rating’**
  - Value within the rating column will be used to designate “good” or “bad” reviews

## Initial Dataset Review Counts:

- ~700K total reviews before filtering
- ~50K users
- ~3.5K businesses



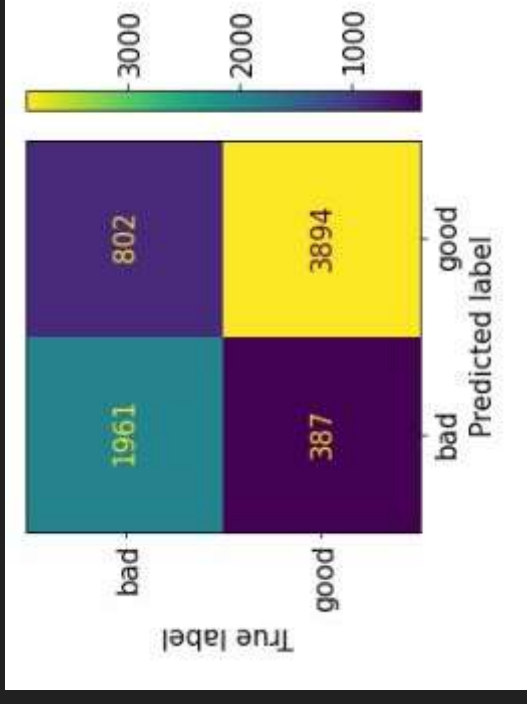
# Modeling

An iterative process of model comparison was used to classify a given restaurant review as “good” or “bad”

**Final Model: Multinomial Naive Bayes, TF-IDF Vectorizer**

## Interpretation of Confusion Matrix:

- 3,894 reviews were correctly predicted as Good
- 1,961 reviews were correctly predicted as Bad
- 802 reviews were incorrectly interpreted as Good, when actual review type was Bad
- 387 reviews were incorrectly interpreted as Bad, when actual review type was good



## Confusion Matrix for Final Model

Model was able to classify a good/bad with an accuracy of 84%

Model is also relatively accurate in the face of testing data (83%)



# Conclusion

Words/phrases that were most impactful for the performance of the predictive model described a restaurant's features/experiences. The following aspects of improvement were identified.

- Food quality (see green words)
- Service quality (see blue words)
- Value perceived by customer (see red words)
- Time to receive food (see yellow words)

Positive Reviews: Words of Highest Importance

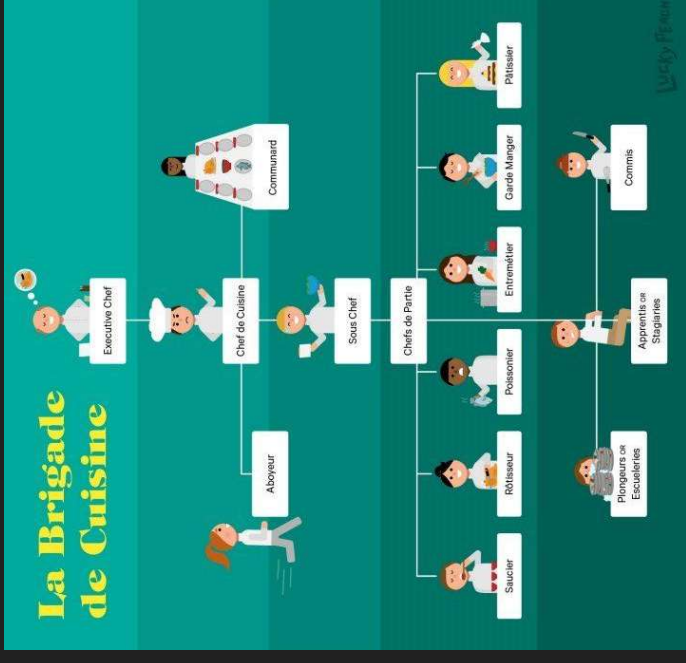
['adobada', 'bouche', 'owner chef', 'masala', 'wine gravy', 'pomo', 'arepas', 'combo pan', 'tikka', 'oggie', 'pmakase', 'combo pan roast', 'postino', 'jjanga', 'lobster mash potato', 'lobster mash', 'limoncello', 'huli', 'coconut bark', 'huli chicken', 'tikka masala', 'huli huli', 'detroit', 'huli huli chicken', 'pan roast', 'tonkatsu', 'cibo', 'red wine gravy', 'dw', 'giada', 'arepa', 'komex', 'staff amaze', 'butter cake', 'raku', 'rollin smoke', 'snoh', 'skinnyfats', 'rollin', 'goodwich', 'soho', 'chicken skin', 'robuchon', 'ike', 'juan', 'michigan', 'east valley', 'kalua', 'kalua pig', 'al pastor', 'yuzu', 'rival', 'mastros', 'best taco', 'pb', 'oyster bar', 'bone ribeye', 'best thai', 'coney', 'sweet savory', 'black cod', 'delicate', 'mr mama', 'banana nut', 'tataki', 'carne asada fry', 'ingredient fresh', 'albacore', 'ayce sashi', 'cornish', 'pizzeria', 'cornish pasty', 'cauliflower', 'taste menu', 'cream corn', 'style pizza', 'empanada', 'sorbet', 'bread butter', 'el gordo', 'take first', 'taco el', 'taco el gordo', 'octopus', 'family run', 'katsu', 'bark', 'owner come', 'british', 'healthy side', 'chicken katsu', 'sf', 'shawarma', 'mastro', 'sandwich ever', 'restaurant town', 'prix', 'place real', 'tztziki', 'spiciness']

Negative Reviews: Words of Highest Importance

['resort fee', 'frozen hot', 'serendipity', 'golden corral', 'dance floor', 'frozen hot chocolate', 'venetian', 'bull', 'hooter', 'dim sum', 'lind', 'mimi', 'corral', 'sport bar', 'sugar factory', 'speak manager', 'waste time money', 'rudely', 'service terrible', 'beer cold', 'service lack', 'hot chocolate', 'tvs', 'waiter take', 'go downhill', 'gratuity', 'disaster', 'drink special', 'front desk', 'trap', 'crab legs', 'heat lamp', 'ugh', 'save money', 'olive garden', 'breakfast buffet', 'bouncer', 'conference', 'downhill', 'watch game', 'full price', 'avoid place', 'tequila', 'service bad', 'pay extra', 'take forever get', 'get refill', 'bring check', 'service horrible', 'monte carlo', 'dj', 'ask manager', 'good either', 'stayed', 'horrible experience', 'cotton', 'cotton candy', 'wing good', 'least minute', 'terribly', 'zero star', 'voucher', 'seat table', 'drinks', 'filthy', 'want drink', 'filipino', 'go bar', 'service okay', 'straw', 'mein', 'waste money', 'good minute', 'service poor', 'frustrate', 'go somewhere else', 'shower', 'draft beer', 'give menu', 'run around', 'say oh', 'chinese restaurant', 'seem care', 'desk', 'good server', 'blonde', 'bring drink', 'get bill', 'service suck', 'broken', 'say sorry', 'waiter say', 'subpar', 'bug', 'hospital', 'drink take', 'watch sport', 'actually really', 'grace', 'worse']

# Recommendations

- Establish a French Brigade system, or some form of improved organizational system
- Establish a happy working culture
- Establish fair front-of-house compensation
- Optimize food spend while maintaining quality
- Incorporate more special features to engage customers





## Next Steps

- More nuanced system of rating, similar to metacritic
- Utilize more filtering options in the recommendation system to mimic Yelp filters
- Create word cloud distributions for each restaurant recommendation for more user context
- Provide additional color to potential restaurant improvements, using additional feature flags provided within Yelp's dataset
- Replicate analysis for different restaurant category types; different kinds of restaurants may require different solutions

# Recommendation System

Appends self-reported user ratings and generates a list of recommendations based on the other users/ratings available in the dataset

- Inputs: user ID, number of restaurants to review, category

```
business_id  name  categories
29764  15FvokE2-k0jPOvUHSrVA  Sushi Catcher  Sushi Bars;Japanese;Restaurants
How do you rate this restaurant/business on a scale of 1-5, press n if you have not been :
2
92481  SP6Zxmt8_nT30-nVnKHYKA  Akaihana Sushi & Grill  \
categories
92481  Sushi Bars;Japanese;Restaurants
How do you rate this restaurant/business on a scale of 1-5, press n if you have not been :
1
35796  6syWU43FK6cXB2957Ga8A  Sekai Sushi  Sushi Bars;Japanese;Restaurants
How do you rate this restaurant/business on a scale of 1-5, press n if you have not been :
2
87088  V3rUBXjLGWniPmPQ0zRhiv  Makino Restaurant  \
categories
87088  Sushi Bars;Japanese;Restaurants
How do you rate this restaurant/business on a scale of 1-5, press n if you have not been :
3
86546  081FR_ruR9Gk3Q6sakt_g  Sushi Creek  Sushi Bars;Restaurants
How do you rate this restaurant/business on a scale of 1-5, press n if you have not been :
4
Out[45]:
```

| business_id             | name                            | categories                      | rating   |
|-------------------------|---------------------------------|---------------------------------|----------|
| 6X9iyuM2XocGT4q9v6cA    | Sushi Bars;Japanese Restaurants | JJANGA Japanese Restaurant      | 4.492273 |
| OWbPIJuz2oUJ5K3DzcHKBtg | Sushi Bars;Japanese Restaurants | Soho Japanese Restaurant        | 4.491927 |
| HpaYCM_MCau7ZLLXc6SA    | Tapas/Small Plates              | Sushi Bars;Japanese Restaurants | 4.489452 |
| sNBqulTAv3ibUWkzSUITpw  | Sushi Bars;Japanese Restaurants | Yonaka Modern Japanese          | 4.489281 |
| Pgp3gpOQaJdyjqC9AOx6g   | Sushi Bars;Restaurants          | Sakana                          | 4.489053 |
| bpGFrX8muk0GxA13Oea3Q   | Sushi Bars;Japanese Restaurants | I Love Sushi                    | 4.489053 |
| u6EUOXO5FqjvzLIHD21bA   | Sushi Bars;Japanese Restaurants | Yummy Grill & Sushi             | 4.485786 |
| i5Tq_Lzm3osbY_uEJ79N1g  | Seafood                         | Hatumi Sushi                    | 4.485590 |
| QrNRFISxmCo4pzVLahA_bw  | Seafood                         | Tokyo Boys                      | 4.480236 |
| AY6ub5UCAngroJ6uJG7ILQ  | Steakhouse                      | Kanji Steak & Sushi             | 4.479729 |
|                         | Sushi Bars;Japanese Restaurants | Kabudo                          | 4.475067 |

## Recommendation System and Output

# Thank You!

Email: [noahiz917@gmail.com](mailto:noahiz917@gmail.com)

Github: noahiz917

LinkedIn: [linkedin.com/in/noah-john-hizon/](https://www.linkedin.com/in/noah-john-hizon/)