Hotel Satisfaction Reviews

Project Report

Shaylee Davis

December 7, 2021

Proposal

Problem Statement

When users create reviews on Trip Advisor there are multiple factors that go into their ranking. People prioritize different factors. Whether the hotel is close to restaurants/shopping, does it have breakfast, does it have a restaurant, how far it is from the airport. In this project we will sort through the user reviews and use natural language processing to find what is most prioritized in our reviews in order to advise hotel management and ownership on what should be their primary concerns in day to day operations.

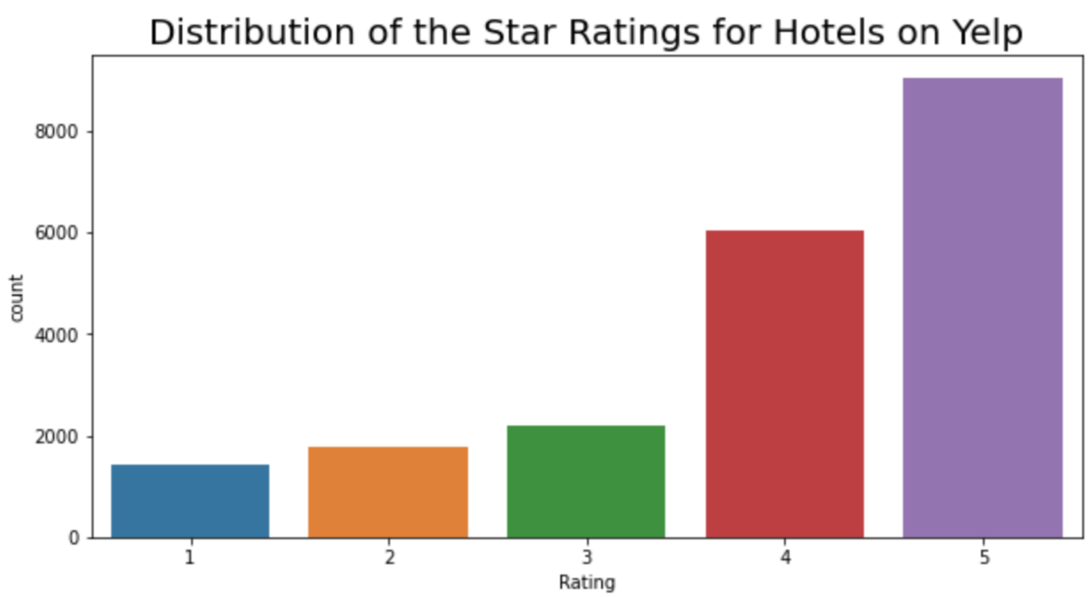
Examination of Trip Advisor Hotel Reviews

Data Wrangling

Data was found from one source Trip Advisor via Kaggle. Data had no missing values. There was no data manipulation required outside of using automated processes such as count vectorizer.

Exploratory Data Analysis

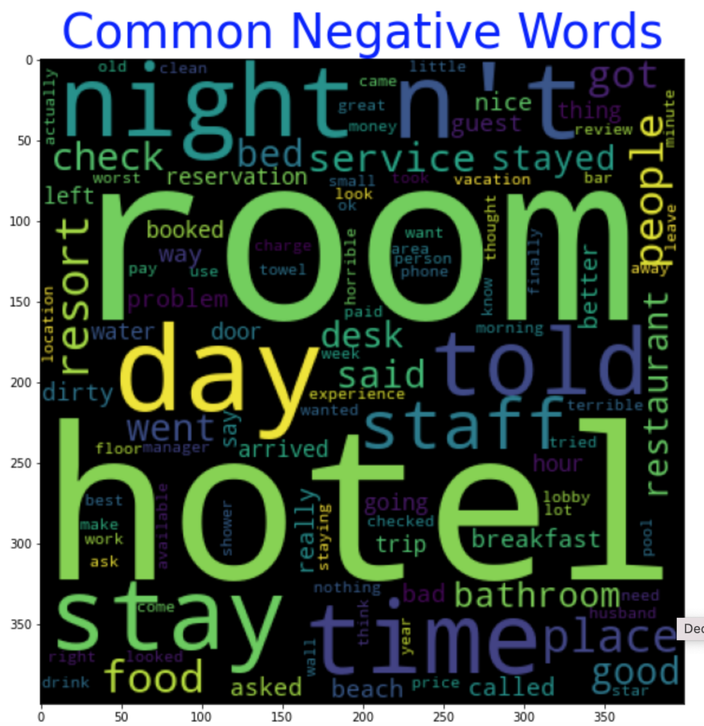
First we searched for the distribution of our data:



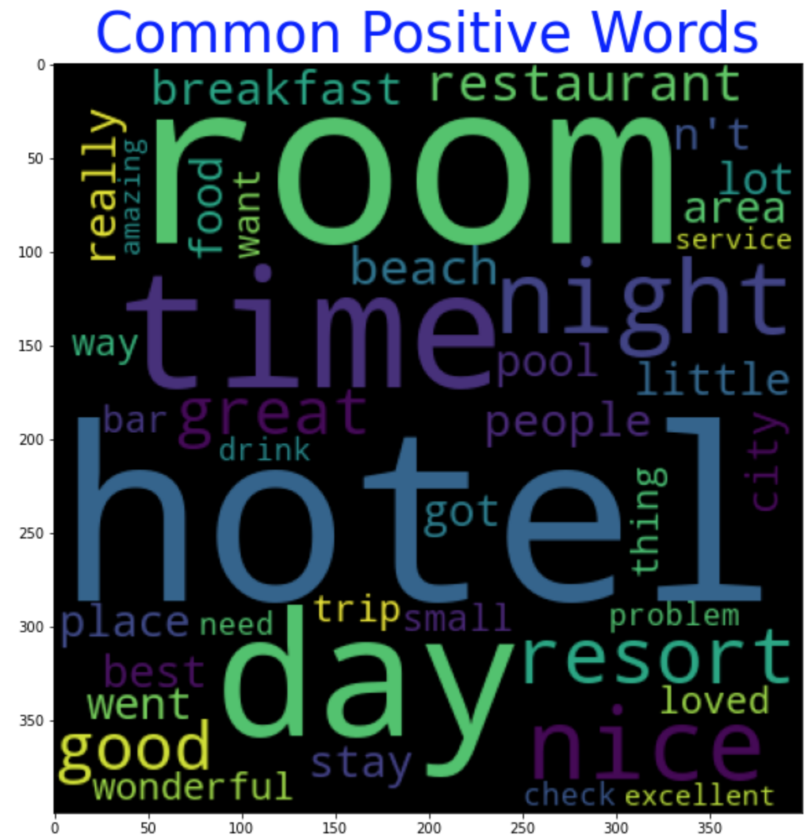
beach (10061), day (9967), and breakfast (9737). Right off the bat without machine learning we can't glean a lot of information froWe found that the people leaving the most reviews were satisfied with their service. Out of the 20,491 reviews we examined, customers left 15,093 excellent reviews (4-5 stars). Which is 74% of our data.

We then examined common words hotel (49814), room (35331), not (31709), great (21475), good (17412), staff (16633), stay (15411), did (14006), just (12667), nice

(12643), rooms (12401), no (11846), location (11351), stayed (10500), service (10367), night (10151), time (10120). We know the words hotel/room/stay/location are going to be frequent, but this doesn't give us any workable insight.

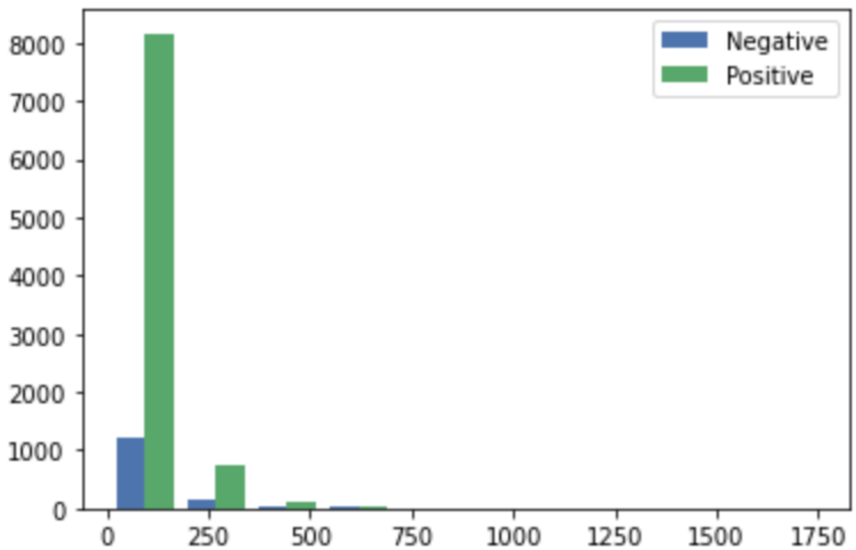


Outside of common words that apply to all hotel reviews (room/hotel). We can see in our negative reviews they spoke frequently of service (835), place(649), and resort (633)



Outside of common words that apply to all hotel reviews (room/hotel). We can see in our positive reviews they spoke frequently of location (4,962), breakfast (4,370), service (4300), clean (3,935), and excellent (3792)

We then looked at the length of our words. When graphed together you can see they both skew towards more words. We had a higher quantity of positive reviews, but it appears the average amount of words is about the same. Positive reviews do have a slightly higher mean of 112 versus negative's 94.




```
count    9054.000000
mean       93.963773
std        90.977999
min         7.000000
25%        44.000000
50%        69.000000
75%       111.000000
max      1755.000000
```

Preprocessing and Feature Engineering

We used a count vectorizer and for our first model only used the most extreme reviews (1 & 5). After that we built a Multinomial Naive Bayes model.

Machine Learning Model

Based on our Naive Bayes Model with 98% accuracy we are able to examine which words held the most weight in determining the rating. In order to find interesting data based on the language used. We examine how many times each token appears in 1 or 5 star reviews. Then we examine a ratio of how often the words occurred in 5 star reviews as compared to 1 star reviews. We looked at the top and bottom 10 words within that ratio.

token	one_star	five_star	five_star_ratio
unhelpful	0.041865	0.000441	0.010530
disgrace	0.013321	0.000147	0.011032
unprofessional	0.011418	0.000147	0.012870
rotten	0.011418	0.000147	0.012870
stunk	0.011418	0.000147	0.012870
reeked	0.011418	0.000147	0.012870
disgusting	0.053283	0.000735	0.013790
urine	0.010466	0.000147	0.014040
incompetent	0.010466	0.000147	0.014040
vomiting	0.010466	0.000147	0.014040

token	one_star	five_star	five_star_ratio
spotless	0.000951	0.041146	43.244673
loved	0.013321	0.225569	16.933820
delightful	0.000951	0.016018	16.834533
brilliant	0.001903	0.031301	16.448420
gem	0.001903	0.031154	16.371198
pricey	0.001903	0.029537	15.521749
eiffel	0.000951	0.014695	15.444526
traditional	0.000951	0.014548	15.290081
wharf	0.000951	0.014254	14.981190
library	0.000951	0.014254	14.981190

We can see the common words for positive reviews were spotless, loved, delightful, brilliant, gem, pricey, eiffel, traditional, library, and wharf. This leads us to see that while cleanliness is one of the top predictors of if a customer will love their hotel the location is also very important. We can see the common words for negative reviews were unhelpful, disgrace, reeked, unprofessional, rotten, stunk, disgusting, incompetent, urine, and vomiting. This leads us to see again cleanliness is one of the top predictors if a customer will love their hotel. The second highest factor in a negative review is staff that is not friendly and accomodating.

Key Takeaways and Future Research

We found that often the people most likely to leave reviews were those with positive experiences. We also found that most people even used more words when speaking of their positive experiences.

When it came to determining whether a review would be positive or negative we found words commonly used for describing the cleanliness of a room (spotless/disgusting) were often used. Another trigger for negative reviews was negative experiences with staff.

For future research I'd like to see if other review sites Yelp/Google have the same ratio of more positive reviews compared to negative. With more data we can increase accuracy of our machine learning models when it comes to predicting ratings 1, 2, 3, 4, and 5.

Final Summary

Based on this we recommend that management prioritizes cleanliness. For ownership location also matters greatly. Other items such as shuttles, continental breakfast, did not seem to matter as greatly. We also recommend that front desk is always staffed and that they are well trained on courteous assistance and de escalation techniques.

Data Sources

Trip Advisor Hotel Reviews

Source: Kaggle

https://www.kaggle.com/andrewmvd/trip-advisor-hotel-reviews

Model Metrics

Final Model: 98% Accuracy