

Recommendations of Airbnb Listings using NLP

Project created by Aisulu Omar
Email: aisuluomar123@gmail.com

WHY?

Destination: Seattle

Do you use Airbnb a lot?

Do you wish it was more personalized?

Did you stay at perfect Airbnb and you
wish you could find more places?



Goal #1

You are planning to go to Seattle and you have no idea what is where.

- You like bars and clubs
- You like quiet places

The screenshot shows the Airbnb search results for "Seattle · Stays". At the top, there are filters for Dates, Guests, Work trip, Type of place, Price, Instant Book, and More filters. A message indicates over 400,000 guest reviews with an average of 4.8 out of 5 stars. Below this, a section titled "Places to stay" lists categories: Entire homes, Airbnb Plus, Unique stays, Bed and breakfasts, and Boutique hotels, each with a representative thumbnail image. A map of Seattle and surrounding areas displays price ranges for different neighborhoods, with labels such as \$59, \$75, \$85, \$92, \$94, \$76, \$89, \$98, \$119, \$78, \$40, and \$20. At the bottom, a specific listing for "ENTIRE GUESTHOUSE Cottage in the Heart of the City" is shown with a thumbnail image of a wooden interior.

Goal #1

I decided to create more personalized recommendations with specific patterns for everyone.

Dataset

I was lucky. The dataset was provided by Airbnb on Kaggle.

I used two datasets:

- First one included all the info on Airbnb listings in Seattle.
- Second one provided reviews of users.

3818 listings

84831 reviews

The process #1

By using NLP toolkits, I was able to turn content of listing and the search query into vectors.

Cosine similarity allowed me to find top 5 listings with the most similar patterns based on the search query.

The outcome #1

New Tab | Project3 _add | Home Page - S | Reviews sentinel | Project_4 - Jup | [rejected] failed | Seattle - Stays | obscure-depths | Heroku App to +

obscure-depths-25051.herokuapp.com

This is app that provide 5 Airbnb recommendations in Seattle based on your search

I am using NLP and cosine similarity



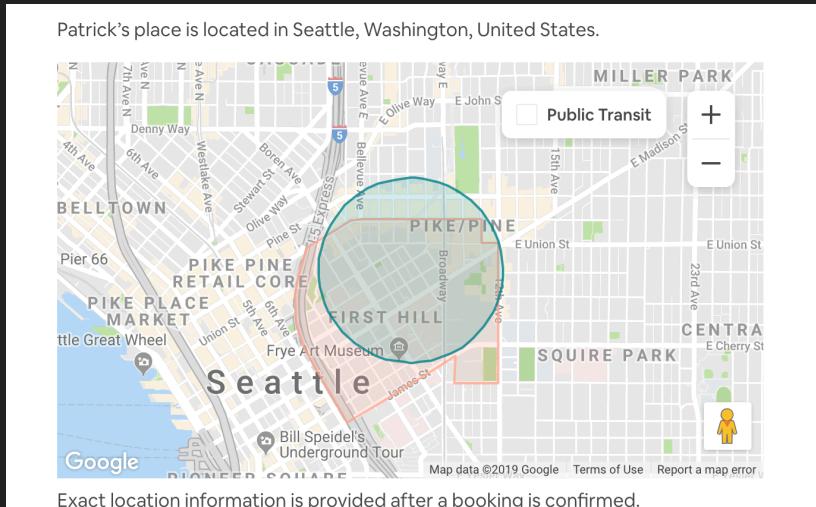
Follow this link

DWwdBcLVAAA...jpg ^ Show All X

A screenshot of a web browser window showing a recommendation page for Airbnb in Seattle. The main heading is "This is app that provide 5 Airbnb recommendations in Seattle based on your search". Below it, a subtext says "I am using NLP and cosine similarity". A large image shows three people from behind, looking out over a body of water at the Seattle city skyline. At the bottom left, there's a link "Follow this link". The browser's address bar shows the URL "obscure-depths-25051.herokuapp.com". The browser interface includes a tab bar with various open tabs and a navigation bar with back, forward, and search icons.

More about the outcome #1

Listing 1 location:



Listing 2 description:

We are renting out a private room in Capitol Hill with a full airbed while our fourth roommate is out of the country. With mountain and lake views and easy access to a ton of parks and bars, this is a low-key spot that's a great place to crash! You'll have your own room, and access to the washer, dryer, kitchen, bathroom (there's only one), and living room if you so please. We all work during the day, so our interaction may very well be limited, but I or my roommates would be happy to give you recommendations on parks, bars, things to do, bus routes, etc.

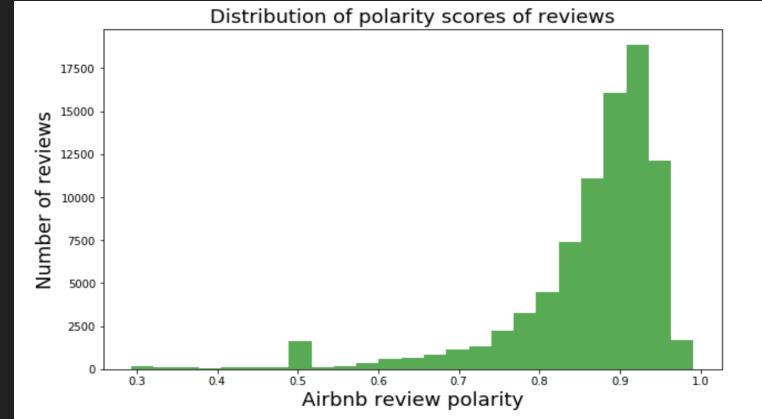
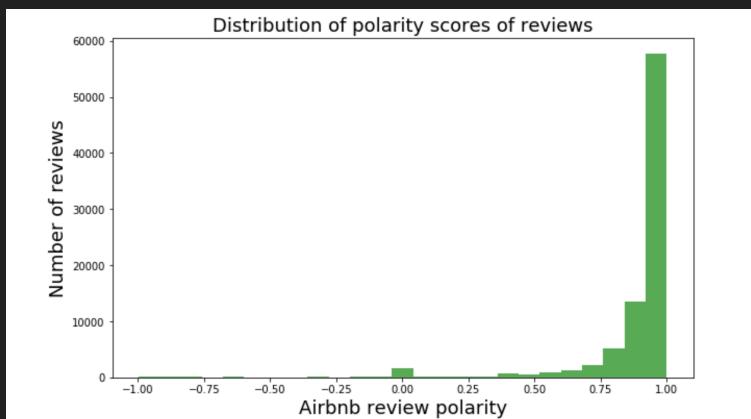
Goal #2

You are an Airbnb user and you had a perfect experience, you left a great review for the listing.

I created 3 recommendations for every good review.

Process #2

NLTK has a package Sentiment Intensity Analyzer that allowed me to identify polarity for every comment.



Process #2

- I created a recommender engine using Surprise (Simple Python Recommendation System Engine).

cross validated RMSE = 0.103

- I wrote a function that recommends up to three listings for every user.

UserId	Recommended Listing1	Recommended Listing2	Recommended Listing3
5259194	17477014,4.934	14453508,4.902	3202492,4.89809
119103	2772019,4.9995	20332653,4.997	5008344,4.98449
4241862	18655414,5	21788547,4.859	11635199,4.859
3487824	14011827,4.890	38765533,4.813	18951949,4.622
1984962	193349,4.96865	18109567,4.968	7496011,4.96869
9531	28371203,4.780	3509427,4.7805	8240057,4.78059
5021081	47728985,4.289	11525944,4.289	38369338,4.289
1539476	10192891,5	149191,4.95411	21052081,4.954
973859	17057653,4.979	11016487,4.895	24913458,4.895
1914051	1893319,4.9272	40654229,4.901	27435868,4.866
443736	11937279,4.764	1593262,4.7649	5919992,4.76498
1097520	11077274,4.937	6475379,4.8321	24270679,4.832
8101251	3512746,4.8161	11230629,4.793	null
1633025	6426992,5	3831403,4.9721	4274504,4.95259
4863533	7704843,4.8151	37367075,4.815	null

Let's check

Recommendations for Lauren UserId 7853322



Cottage in the Heart of Ballard
Seattle

Brad / Liz

Entire guesthouse
2 guests • 1bedroom • 1bed • 1bath

Sparkling clean
10 recent guests said this place was sparkling clean.

Brad / Liz is a Superhost
Superhosts are experienced, highly rated hosts who are committed to providing great stays for guests.

\$75 per night
 550

Dates
Check-in → Checkout

Guests
1 guest

Great place, I love the coziness and natural atmosphere of the house. I enjoyed my stay.

Let's check

Recommendations for Lauren UserId 7853322



Urban Cottage-Own Bath-near UW & Park City View



Regina

This place is getting a lot of attention.
It's been viewed 500+ times in the past week.



Entire house
2 guests 1 bedroom 1 bed 1 bath

Sparkling clean
12 recent guests said this place was sparkling clean.

\$75 per night
★★★★★ 343

Reserve



Tiny House in Seattle
Seattle



Tipsy The Tiny

This place is getting a lot of attention.
It's been viewed 500+ times in the past week.



Tiny house
3 guests 1 bedroom 1 bed 1 bath

Sparkling clean
15 recent guests said this place was sparkling clean.

\$119 per night
★★★★★ 446

Reserve



SECOND STREET SUITE -- "The Roost"
Langley



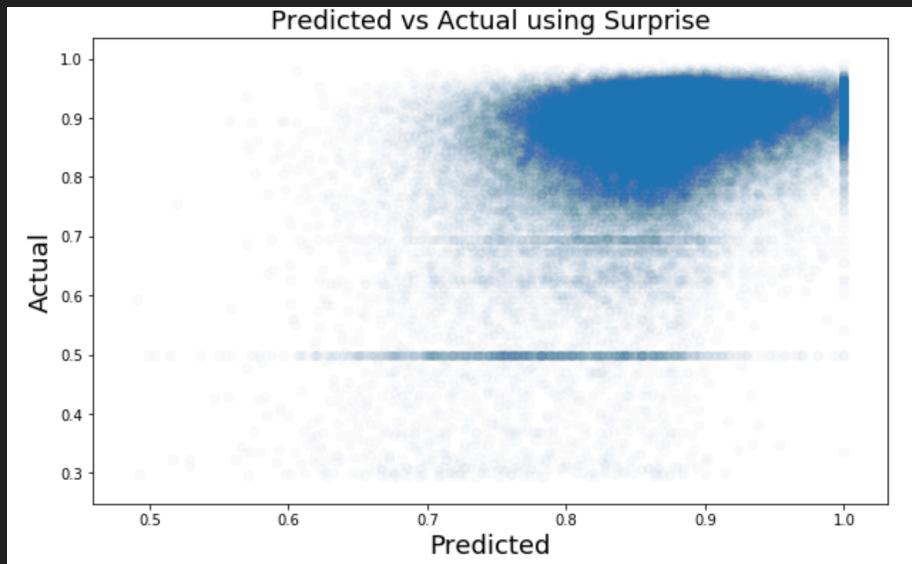
GENE AND TAMMI

\$179 per night
★★★★★ 157

Dates
Check-in → Checkout
Guests
1 guest

Validate the process

After I validated results, by creating a plot of predicted versus actual datapoints.



Summary

Part I of this project is successfully implemented through Flask App in Heroku.

Part II exists in my Jupiter notebook and for improving my model, I need more comments with bad experience.

Thank you!