Airbrb: Predicting Loyalty

Kaciny Calixte
SUNY Old Westbury
kacinycalixte@gmail.com

Jacqueline Curran
Manhattan College
jcurren.student@manhattan.edu

Louise Y. S. Lai
New York University
louiselai88@gmail.com

Erica Ram
Adelphi University
eram1294@gmail.com

INTRODUCTION

Airbnb has two million rooms available - nearly three times more than that of the largest hotel chain in the world.

Airbnb's customers can be broken down into two groups: guests and property hosts. How can we predict if a guest will continue booking rooms through Airbnb, and how can we predict if a host will continue listing out their property on Airbnb? Unlike traditional hotels, Airbnb relies on one group of its customers (hosts) to generate the supply of rooms for its guests. Thus, predicting customer loyalty, both on the host side and guest side, is arguably more important to Airbnb than to traditional hotels. In this paper, we try and capture this loyalty to Airbnb.

The data reveals that *guests* who stay with highly rated hosts are far more likely to return to Airbnb. For *hosts*, the best predictors of whether or not they will continue to list through Airbnb is the recency and frequency of their listing.

DATA & METHODS

Our work focused on two datasets collected by the InsideAirbnb website called "listings" and "reviews".

The listings dataset contains information on the properties being listed on Airbnb and the hosts associated with these properties. The reviews dataset includes details on reviewers, the properties they have booked, and the reviews they have submitted. Both of these datasets were scraped from the official Airbnb website on a monthly basis ranging from January 2015 to July 2016. The listings dataset was used to gather host information, such as ID number, start date, property details, and superhost status. The reviews dataset was used to gather guest information such as ID number, review date, review text and the listing ID it is associated with.

It is important to note that all datasets are organized by city, spanning the world. However, for host analysis we focused on New York City because we made the assumption that the majority of hosts will tend to list properties within the same city. Guests, on the other hand, tend to travel to different cities. Guest analysis was broadened to incorporate the entire United States, in order to obtain a more complete picture of guest patterns and behaviors.

All hosts and guests that were active in 2015 were compiled. From here, we extracted/calculated host and guest attributes/features for 2015 and 2016. The 2015 attributes were our predictor features while the 2016 attributes were our target variables. The target variables were meant to signal whether a host or guest has returned to Airbnb in 2016.

Host predictor features were broken up into categories which included price, reviews, date, place, amenities, location, verification, recency/frequency and interplay. Recency/frequency pertains to host and listing activity. Interplay refers to the interaction between hosts and guests. Our host target variables were whether their listing persisted into 2016 and if it continued to receive reviews in 2016.

Guest predictor features were broken up into categories which included words, recency/frequency and interplay. Recency/frequency, in this case, pertained to guest and reviewer activity. Our guest target variables were whether or not a guest has left a review in 2016. It must be noted that reviews was used as a proxy for bookings or stays. We assumed that the majority of guests will leave a review, indicating that they have booked that listing recently.

After organizing our predictor features and target variables into a Guest dataset and a Host dataset, we split both datasets into train and test sets. We created models using decision trees to find patterns in the given data to predict on unseen data. While training our models, we performed cross validation before testing on our test sets.

RESULTS

Measuring Host Loyalty

In order to quantify host loyalty, we looked at two metrics: listed in 2016 and reviewed in 2016. To be considered listed in 2016, a listing must have made at least one occurrence in 2016 on the Airbnb and website. Similarly, to be considered reviewed in 2016, a listing must have had at least one review in 2016. Both listings and reviews in 2016 are useful proxies for measuring host loyalty; however, they have their limitations. For listed in 2016, many listings appear on the website for the entirety of the year, but that is not necessarily indicative of whether or not people actually rented during this time. For reviewed in 2016, not every guest leaves a review so this metric fails to fully capture actual stays. In addition, reviews can be entered for up to two weeks after a stay, so for those guests who lodged in an Airbnb at the end of 2015, they may have reviewed within the first two weeks of 2016. Therefore, there may be some spillover of number of reviews in 2016. Despite these limitations, reviewed in 2016 is more indicative of actual stays and thus the more useful metric for measuring host loyalty.

Once the target was defined, the listings were divided into different cohorts: some defined by time, some defined by length of time listed on Airbnb. The results from all the various cohorts were comparable, so the results presented in the following segment will be reflective of the January Cohort. Each of these listings appeared on the Airbnb website in 2015 for the first time in January.

Two factors that are predictive of loyalty across industries are: how frequent and how recent a person used a platform, known as recency/frequency (Fader et al.). For example, in the car rental market, most of the business comes from the most frequent and most recent renters. Based on this empirical evidence, the baseline tree created was composed of recency/frequency features, as seen in Figure 1. The area under the curve (AUC) for this tree was 0.9. AUC is an evaluation metric for binary classification problems. An AUC of 0.5 is random, and 1 is perfect.

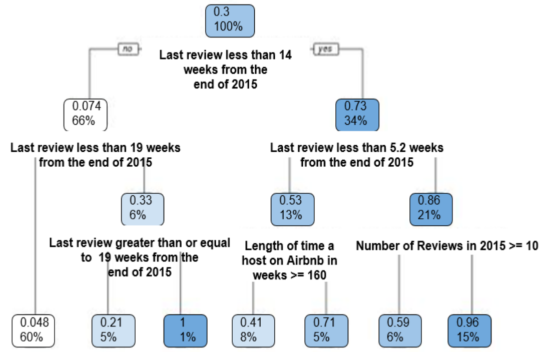

Figure 1: Recency/Frequency Host Tree For January Cohort

After establishing a baseline, different subsets of features were analyzed to see if these features, alone or combined, could perform better than recency/frequency. Although these decision trees did not outperform the baseline recency/frequency model, these features alone were able to perform well in gauging host loyalty. For example, using the words from the review text, resulted in an AUC of 0.79. Similarly, the interplay features resulted in an AUC of 0.86. These AUC values are lower than the recency/frequency model, but shed light on the importance of individual feature subsets.

By combining subsets, such as recency/frequency, and amenities, interesting correlations were ascertained. For example, if a listing has a hair dryer, then that host is likely to be reviewed in 2016 and thus loyal. Although these results do not suggest causation, there appears to be an interesting relationship between features.

In terms of host loyalty, those who are listed most often and most recently, are more likely to be loyal. Additionally, other features, in isolation, are predictive of host loyalty. By combining recency/frequency features and other subsets, the AUC slightly improves and provides interesting insights.

Measuring Guest Loyalty

Based on the data we have, the best metric we have to measure guest loyalty is whether the guest leaves a review in 2016. By having reviews in 2015 and also in 2016, we can see that the guest has returned to Airbnb. However, while using reviews as a proxy for stays is the best metric we have, there are some limitations. It’s more difficult to gauge guests staying again by their reviews because while every host has a listing, guests are not required to leave a review after every stay. Additionally, the exploratory analysis showed that looking at reviews for only New York was insufficient; guests who stay in New York may also stay elsewhere, so in the interest of having more data for modeling, a dataset of reviewer data for the entire United States was used.

Like with the host data, we cohorted the reviewer data to look at reviewers with similar amounts of experience with Airbnb. We looked at cohorts based on the length of time they've been guests and how many reviews they have left. Since similar results were seen between all cohorts, the results presented are of guests who started in January of 2015 and guests who had more than five reviews in 2015.

Our initial tree for the January 2015 cohort was modeled using recency/frequency features (Figure 3). An AUC of 0.65 was achieved, which is lower than modeling host loyalty using recency/frequency, suggesting that gauging guest loyalty is not only different from host loyalty but is also more difficult. Potential reasons for the need for different approaches could be that there are much fewer features in the reviewer data, and also that there is not enough information per guest.

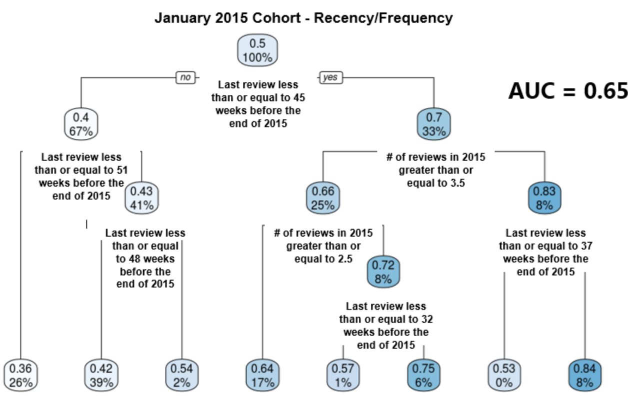

Figure 2: Recency/Frequency Guest Tree For January Cohort

Other subsets of features were introduced, consisting of the most used words in review text and interplay features of the listing from the guest's last review were analyzed to look for better performance than recency/frequency, resulting in both having AUC values of 0.59. These resulted in lower AUC values individually, however when recency/frequency features and interplay features were combined, the AUC improved to 0.66. In this tree, interplay features suggest that guests whose last stay was with a host with a higher rating score are likely to return.

Approximately 70 percent of guests that left reviews in 2015 had exactly one review in 2015, skewing the dataset. When looking at guests whom we had more information about, specifically that had more than five reviews in 2015, we saw a stronger predictive performance using recency/frequency features, resulting in an AUC value of 0.76.

Guest loyalty is best predicted using recency/frequency features, however the interplay features of the listing from the last review do add value to the model. The prediction ranking increases when looking at guests for whom more data exists.

DISCUSSION

Airbnb faces unique challenges because unlike traditional industries, this sharing-economy platform relies on the return of both their hosts and guests. Our predictive models show that reviews and interaction between hosts and guests is of great importance, as well as recency/frequency. Correlations have been found with features such as amenities, which could be useful to test in future experiments.

The incentive is substantial: Airbnb could potentially boost return-rates of first time guests by providing them with incentives to stay at highly-rated properties. Assuming that the average length of stay is 6.4 nights and the average nightly rate is $130, Airbnb could've increase its revenue by $10.1 million, if it were to increase the percentage of returning guests by just 1% in 2015.

ACKNOWLEDGEMENTS

We thank our mentors Shawndra Hill, Amit Sharma, Sébastien Lahaie, Fernando Diaz, and Christopher Riederer for their guidance on this project as part of the 2016 Microsoft Research Data Science Summer School. We also wish to extend a special thanks to Murray Cox, the founder of InsideAirbnb website, for making the datasets easily accessible to the public.

REFERENCES

Fader, Peter S., and Bruce GS Hardie. "Probability models for customer-base analysis." Journal of interactive marketing 23.1 (2009): 61-69.