BOOSTING HOST SUCCESS IN NEW YORK CITY

Identifying factors that decrease the chances of a listing being reviewed

EXECUTIVE SUMMARY

More than 20% of the Airbnb listings in New York City do not have any reviews, which for most of these listings means that they have never been booked at all. Through statistical analysis we have developed a model that can be used to identify these listings based on other information, such as minimum nights, availability, and host listings count.

Identifying these listings that are unlikely to be reviewed will allow Airbnb to intervene, thereby both boosting bookings in the city and protecting hosts from falling into the cycle of lack of reviews leading to lack of bookings. The model was developed with interpretability in mind, so it is clear which features of the listing have the greatest effect and how.

According to the model and statistical analysis, factors that greatly reduce the likelihood that a listing will receive a review are price (above $200), and minimum nights (above 3). Availability and host listings count also play a role but their impact depends on other factors.

In this report, the performance of the model is evaluated and its predictions visualised.

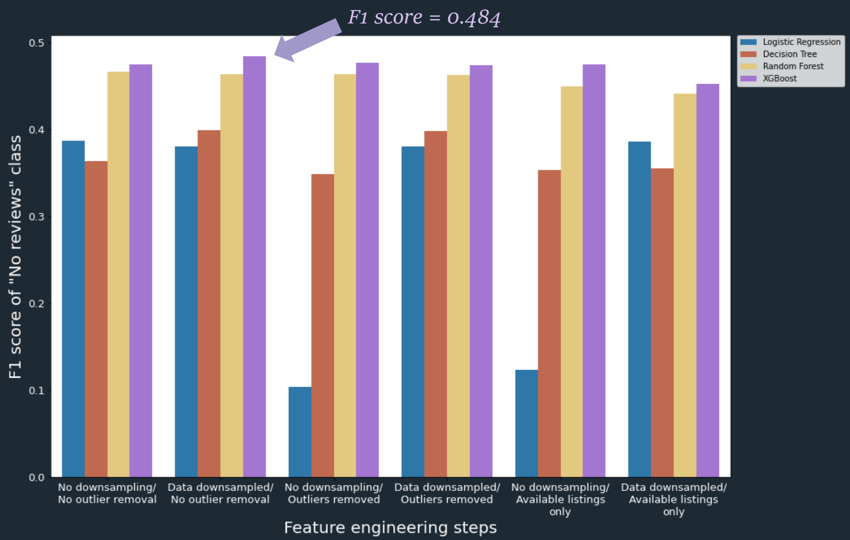


Figure 1 Model performance for 24 trained models. The purple bars show the Xgboost models trained after different feature engineering steps. The F1 score is a measure that combines recall and precision (see 'Performance'). The model with the highest F1 score is indicated.

THE MODEL

The model used to predict which listings would not receive reviews is an XGBoost model. XGBoost stands for eXtreme Gradient Boosting. It is a model that is made up of several hundred simple models, each of which "learns" from the mistakes of previous simple models. As a result, XGBoost has a high accuracy compared to other models tested. As Figure 1 demonstrates, the model was selected from a total of 24 trained models.

PERFORMANCE

The model correctly identifies 61% of listings that will not be reviewed, whilst only misclassifying 27% of listings that will be reviewed. Whilst there is room for improvement (see the 'recommendations' section), carefully placed interventions could lead to more than 5000 listings being booked for the first time (61% of the un-reviewed listings).

Figure 2 is a confusion matrix for the model that demonstrates its performance. It is important to balance the proportion of listings without reviews that it correctly identifies (recall), with the proportion of listings it predicts not to have reviews that actually do not (precision), so as not to inadvertently interfere with successful hosts. This model was selected because it has a low false negative rate.

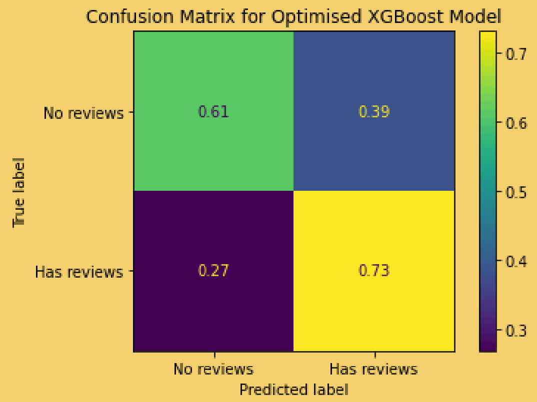


Figure 2 This confusion matrix demonstrates the proportion of "true" values the model predicts to be either a listing that "has reviews" or has "no reviews". For example, of all of the listings without reviews, the model predicted that 61% of them would not have reviews, and 39% of them would have reviews.

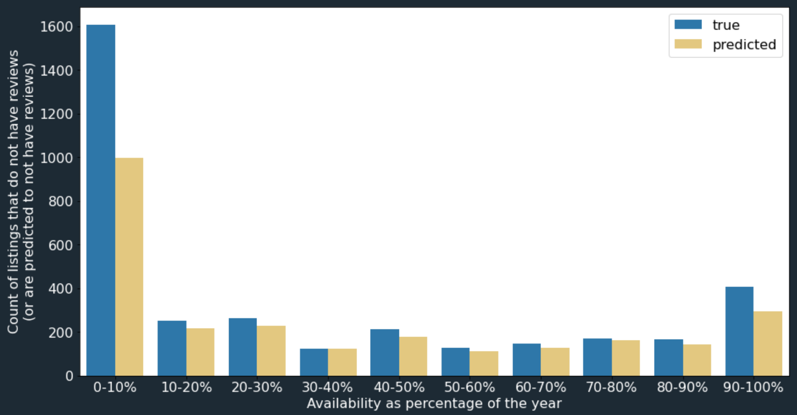


Figure 3 The number of listings that have reviews (blue) and the number of listings that are predicted to have reviews (yellow). For listings with 0 availability, the model underpredicts reviews.

LIMITATIONS

As Figure 3 demonstrates, the model performs poorly when the availability of the listing in question is set to 0 days out of 365. These listings should not be shown to the model. This should not be an issue however, as listings with availability of 0 cannot be booked and therefore *should not* have reviews.

FACTORS THAT PREVENT LISTINGS FROM BEING REVIEWED

The model's predictions are impacted by a number of factors that interact with one another. A summary of these interactions can be seen in Figure 4.

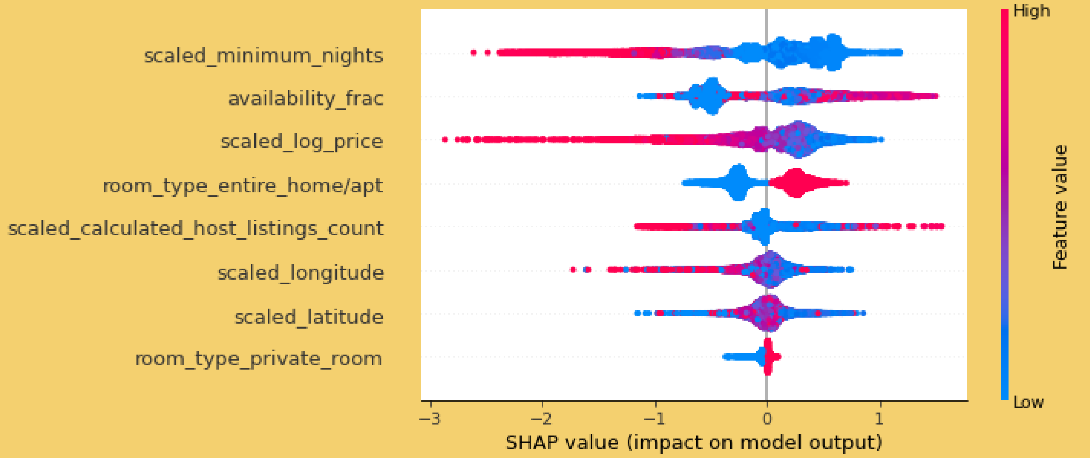


Figure 4 The impact on the model's predictions of its different features. The features provided to the model are ordered from top to bottom by their overall importance to the model (descending). Each dot is a listing in NYC. If it is red, that feature is high for the listing; if it is blue the feature is low. The dot's position on the x-axis shows how it influences the model. The farther to the left, the more it is pushing the model to predict "no reviews". The farther to the right, the more it is pushing the model to predict that it "has reviews".

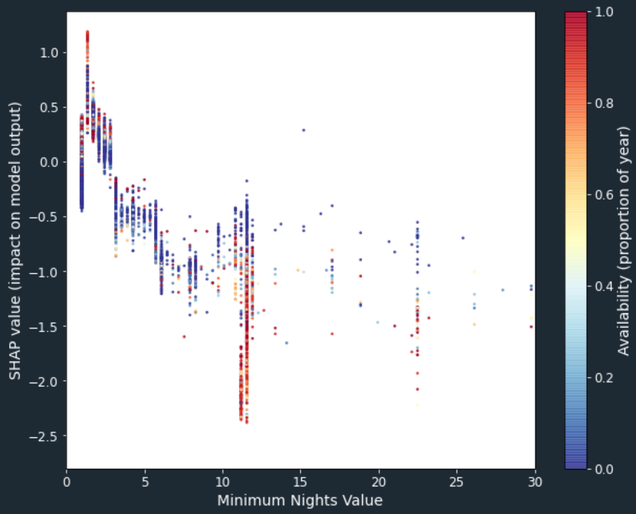


Figure 5 Listings with minimum nights above 3 are less likely to have reviews. Among listings with low minimum nights, those with high availability are more likely to have reviews. Each dot represents a listing, its colour represents its availability, and its vertical position shows its impact on the model. A positive impact makes the model predict "has reviews", whereas a negative impact leads to a "no reviews" prediction.

> "High minimum nights and high price listings are more likely to have no reviews"

MINIMUM NIGHTS, AVAILABILITY, AND PRICE

are the biggest contributors to the model's output, as can be seen in Figure 4. Furthermore, high minimum nights and high price listings are more likely to have no reviews. Figures 5 and 6 give us more detail about these effects. Availability has an unclear effect on the prediction, it seems to depend on other factors. For example, among listings with low minimum nights higher availability listings are more likely to be reviewed, whereas among listings with price > $200 they are more likely to be reviewed.

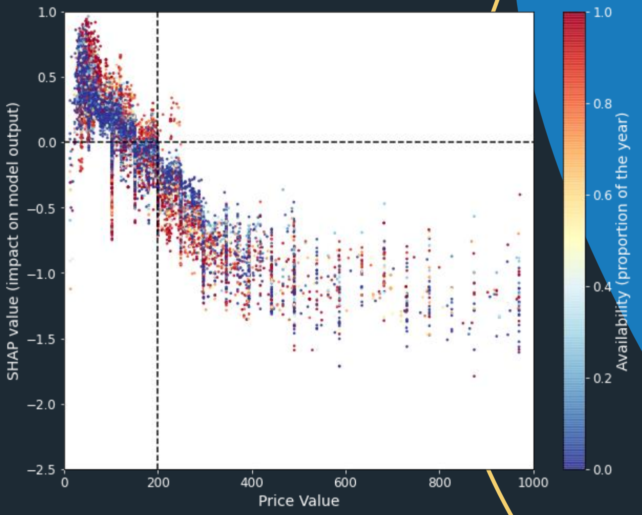


Figure 6 Listings with prices greater than $200 are predicted to have no reviews. Interestingly, among listings with higher prices, those that are more available are less likely to be reviewed. These may be owned by businesses rather than being homes.

ROOM TYPE alone does not appear to have a major impact on whether or not a listing will have reviews, but the other factors known to impact reviews have varying effects on listings of different room types.

For example, listings that are "entire homes" and have reviews have a lower average price than entire homes that do not have reviews, as is suggested in Figure 6. However, private and shared rooms do not show this discrepancy in price between listings with and listings without reviews (Figure 7).

There is a more significant difference in the average minimum nights between listings with and without reviews across all room types (Figure 8).

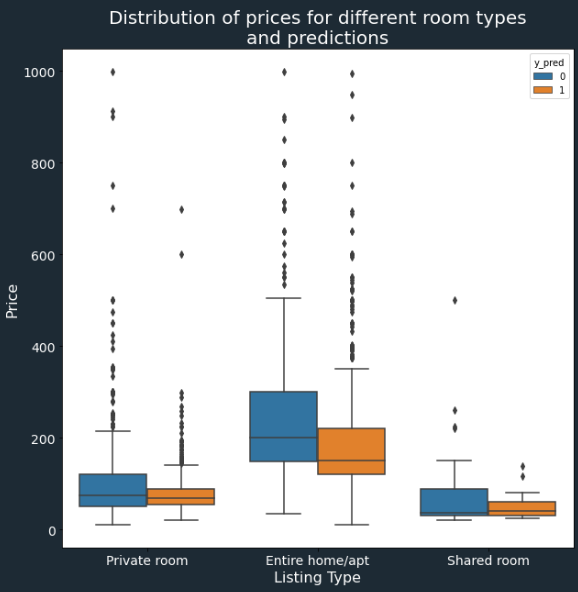
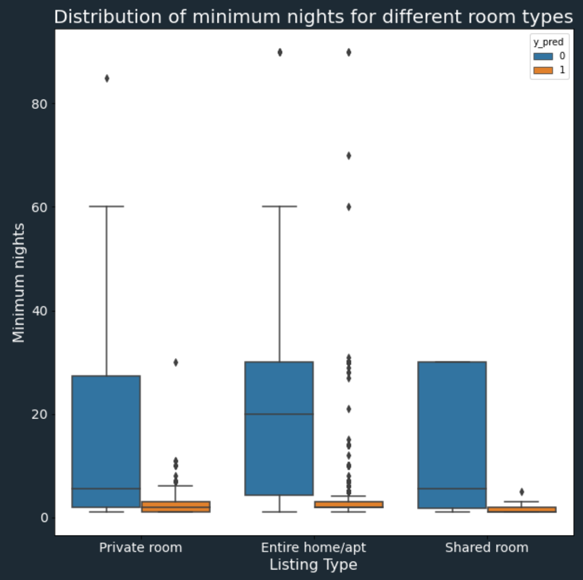


Figure 7 Distribution of prices for listings of different types, grouped by whether or not they are predicted to have reviews (y_pred; 1 means the listing has reviews).



OVERALL, the model tells us that listings with a high number of minimum nights are less likely to be booked and reviewed, especially if they have a high price and are an entire home. Listings with low minimum nights should have a high availability throughout the year in order to be successful, whereas those with high minimum nights perform better when their availability is low. The reason for this is unclear but could be investigated if more data was made available (see below).

Figure 8 Distribution of minimum nights for listings of different types, grouped by whether or not they are predicted to have reviews (y_pred; 1 means the listing has reviews).

RECOMMENDATIONS

Given that listings without reviews are less likely to be booked, Airbnb should intervene when a listing is very unlikely to be reviewed. A message should be sent to hosts when their listings are predicted to have no reviews, along with an explanation of how the model works (the relative importance of the factors listed on the previous page). A warning could also appear on the screen when a host attempts to input a price greater than $200, especially if the listing has a high availability or high minimum nights, to warn them that their chances of success are relatively low.

It would also be advantageous to expand the model's input data to improve its performance. The model's accuracy of 61% could be improved by providing more information about the listings. Furthermore, providing more information would help Airbnb to understand what features of a listing lead to bookings. For example, providing listing amenities and measuring the impact of each amenity on the model's output would demonstrate which amenities are most valuable to prospective guests.

It is likely that the most import factor in whether or not a listing has been reviewed will be its time on Airbnb, or at least, how long it has been available for. Listings that have not been available for long will not have reviews, but for completely different reasons to listings that have been available for several months and not been reviewed.