

AirBnb Booking Prediction

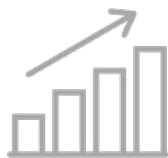
Mid Project Review

Developer : Shreyas Sabnis
QA: Carson Chen



Highlights

- 1) Data Cleansed to eliminate incorrect data points, handle missing values and correct datatypes for training data.
- 2) Exploratory data Analysis performed to get a sense of the distribution of various features and their relative importance in predicting the response variable (Destination Country)
- 3) Interesting findings about the data:
 - Specific destination countries seem to favor younger users, while others favor older users
 - Highest % of bookings were local (within USA)
 - The website sees very little mobile traffic.



Review Progress

Stories completed:

1. Model Development . Data Cleansing (2)
2. Mode Development . Exploratory Data Analysis (4)
3. Model Development . Outlier detection and Management (2)

Stories in Progress:

1. Model Development. Feature Engineering (8)



Demo

Jupyter Notebook – [Cleansing_EDA.ipynb](#)



Lessons Learnt

- The user sessions dataset has user browsing actions that may be useful in generating features to predict the outcome. Include that in analysis during future sprints.
- Age appears to be an important factor, however it has a very large proportion of missing data, so it may not end up being a significant predictor.



Recommendations

Stories to tackle in upcoming sprint:

1. Model Development . Feature Engineering (8) – IN PROGRESS
2. Model Development . Model selection and parameter tuning (8)
3. Model Development . Model Evaluation (2)