

New York City Airbnb Analysis

By Jennifer Ha
August 27, 2021





1

Business Problem



2

Data



3

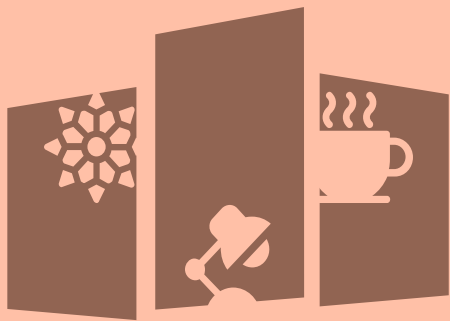
Results



4

Next Steps

Contents



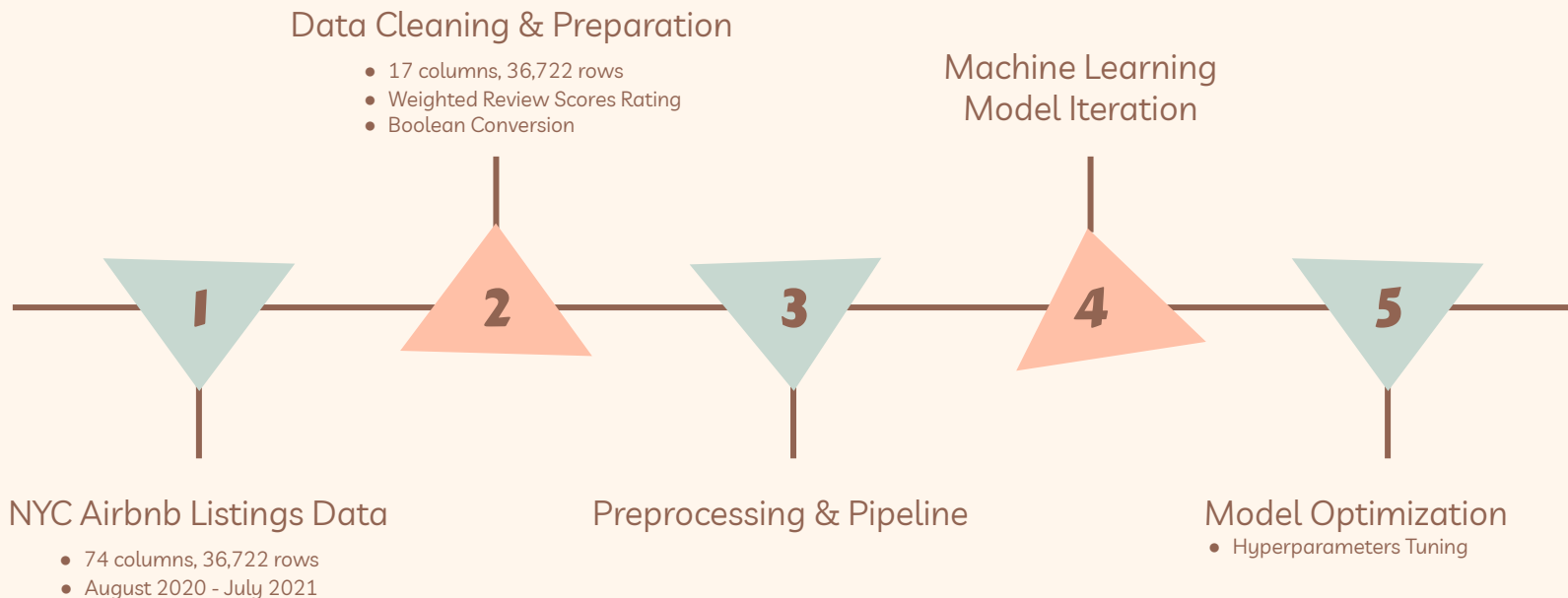
Business Problem

Predict whether a listing is valuable or not

Top 25% of the listings with the highest weighted review scores rating are considered to be valuable listings.

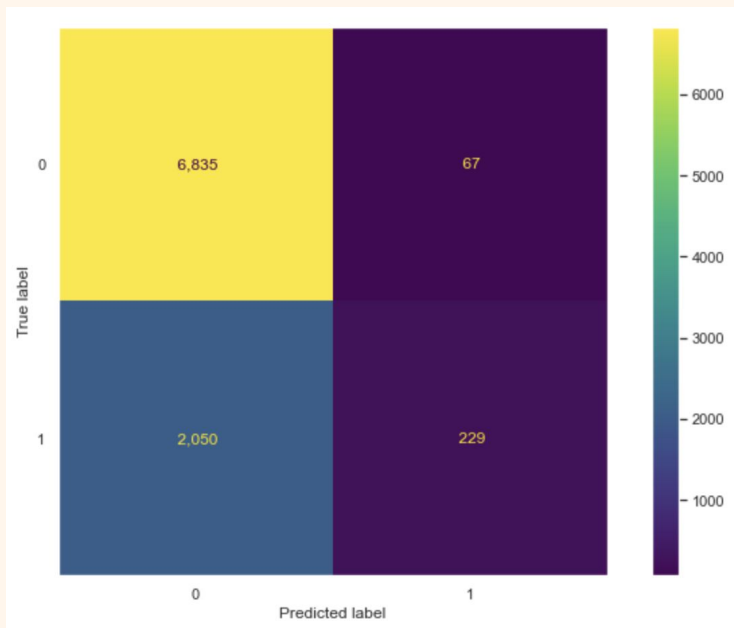
Explore different types of machine learning models and share the best performing model to the Airbnb marketing team.

Data & Methods



Results

- Our winning model: the gradient boosting model with hyperparameters tuning
- increased the precision by 18.5% with almost no sacrifice to the accuracy and ROC-AUC score
- showed highest threshold when deciding whether a listing is valuable or not

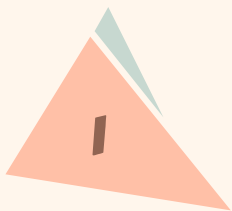


Train Scores

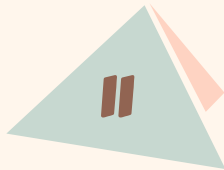
```
-----  
Accuracy: 0.7732471587814531  
Precision: 0.84400826446281  
ROC-AUC: 0.820293822299771  
-----
```

Test Scores

```
-----  
Accuracy: 0.7694150963947283  
Precision: 0.7736486486486487  
ROC-AUC: 0.7875422974866968
```



Understand what features makes a listing valuable



Re-visit over-fitting models and think about how we can reduce the over-fitting



Consider expanding this project to make predictions using global data

Next Steps

Thanks

Email:

jnha1119@gmail.com

Github:

@jennifernha

LinkedIn:

<https://www.linkedin.com/in/jennifernha/>