FLATIRON HOTELS

Predicting Booking Cancellations

Norman Jen

Yamuna Umapathy

BUSINESS PROBLEM

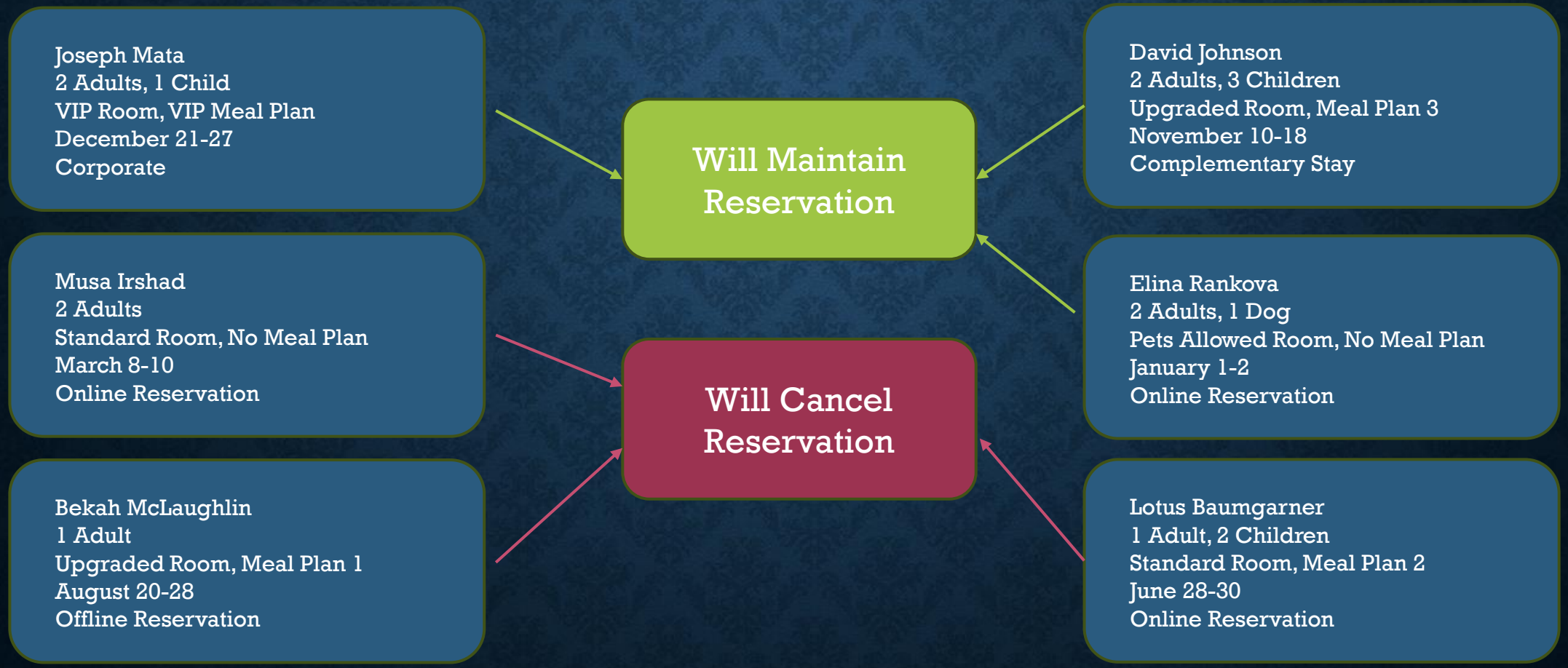
Accepting Cancellations

- Loss of income
- Over-staffing
- Waste of resources
- Turning away guests

Assuming Cancellations

- Upset guests
- Over-booking
- Compensation
- Loss of repeat guest

SOLUTION


Joseph Mata
2 Adults, 1 Child
VIP Room, VIP Meal Plan
December 21-27
Corporate
Musa Irshad
2 Adults
Standard Room, No Meal Plan
March 8-10
Online Reservation
Bekah McLaughlin
1 Adult
Upgraded Room, Meal Plan 1
August 20-28
Offline Reservation
Will Maintain Reservation
Will Cancel Reservation
David Johnson
2 Adults, 3 Children
Upgraded Room, Meal Plan 3
November 10-18
Complementary Stay
Elina Rankova
2 Adults, 1 Dog
Pets Allowed Room, No Meal Plan
January 1-2
Online Reservation
Lotus Baumgarner
1 Adult, 2 Children
Standard Room, Meal Plan 2
June 28-30
Online Reservation

Source:
www.sciencedirect.com

Location – Portugal

THE DATA

Number of Guests ✓

Length of Stay ✓

Lead Time ✓

Repeated Guest ✓

Room Price ✓

Method of Booking ✓

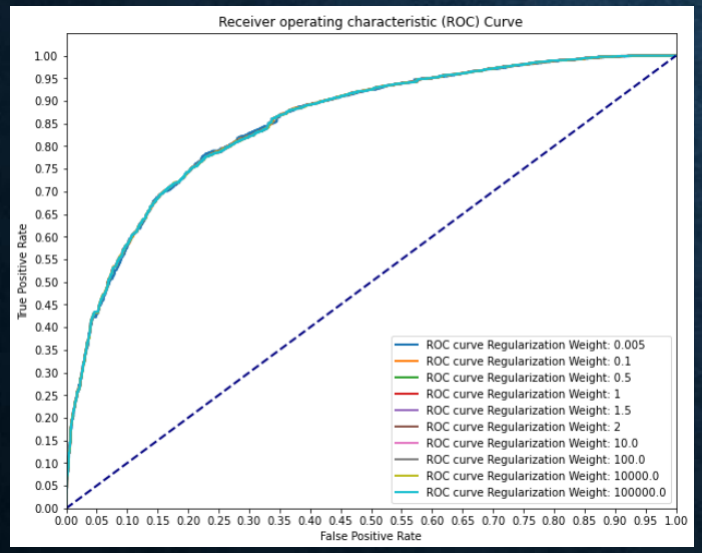
Booking Date ✓

Special Requests ✓

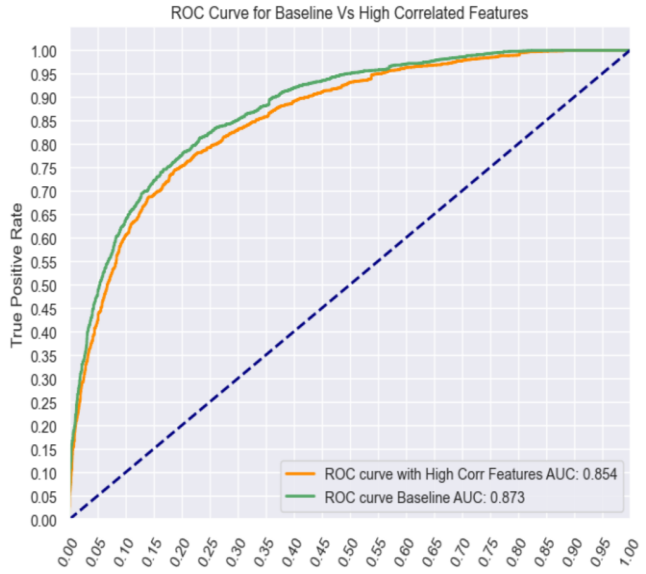
CLASS IMBALANCE

```
booking_status
0                17090
1                 8276
dtype: int64
```

```
1    17090
0    17090
Name: booking_status, dtype: int64
```



- Our dataset has 8,276 cancellations and 17,090 held reservations
- Approximate balance is 67% - 33%
- Used SMOTE to create synthetic data
- Had no significant effect on model's performance



Logistic Regression Using All Features

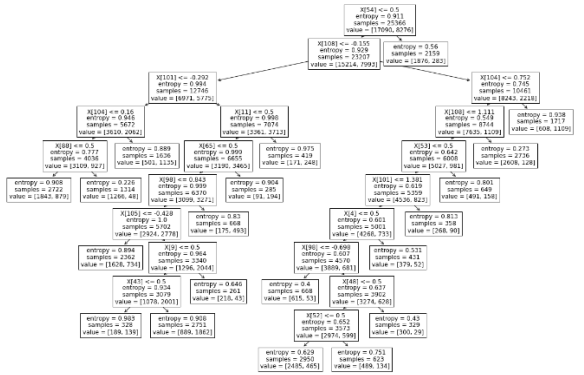
AUC: 0.8722
Train Recall: 0.6474
Test Recall: 0.6213
Train Precision: 0.7478
Test Precision: 0.7551
Train Accuracy: 0.8137
Test Accuracy: 0.8078
Train F1 Score: 0.6940
Test F1 Score: 0.6817

Logistic Regression Using Feature Selection

AUC: 0.8553
Train Recall: 0.6249
Test Recall: 0.6174
Train Precision: 0.7259
Test Precision: 0.7408
Train Accuracy: 0.8006
Test Accuracy: 0.8017
Train F1 Score: 0.6716
Test F1 Score: 0.6735

LOGISTIC REGRESSION

Decision Tree with All Features
AUC: 0.8007
Train Recall: 0.6091
Test Recall: 0.5966
Train Precision: 0.6743
Test Precision: 0.6796
Train Accuracy: 0.7765
Test Accuracy: 0.7732
Train F1 Score: 0.6400
Test F1 Score: 0.6354



Decision Tree with Feature Selection
AUC: 0.8317
Train Recall: 0.5724
Test Recall: 0.5644
Train Precision: 0.6576
Test Precision: 0.6652
Train Accuracy: 0.7633
Test Accuracy: 0.7616
Train F1 Score: 0.6121
Test F1 Score: 0.6107

DECISION TREE

RANDOM FOREST CLASSIFIER

Random Forest Classifier with All Features and No Hyperparameter Tuning
AUC: 0.9359
Train Recall: 0.9879
Test Recall: 0.7873
Train Precision: 0.9960
Test Precision: 0.8907
Train Accuracy: 0.9948
Test Accuracy: 0.8975
Train F1 Score: 0.9919
Test F1 Score: 0.8358

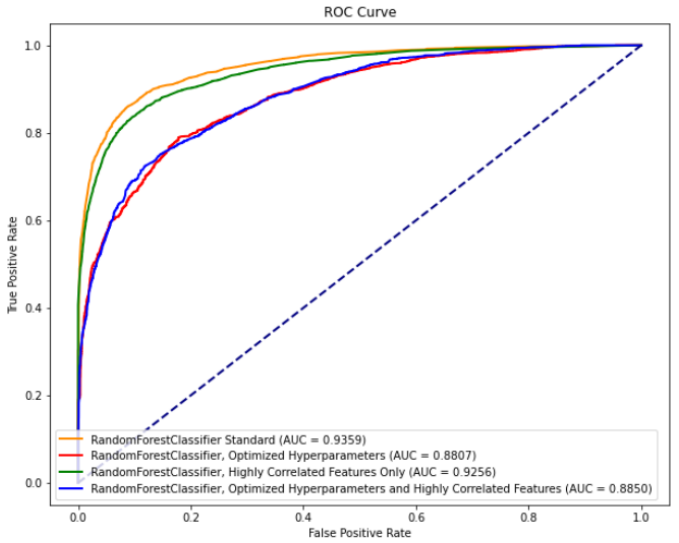
Random Forest Classifier with Feature Selection and No Hyperparamater Tuning
AUC: 0.9256
Train Recall: 0.9848
Test Recall: 0.7868
Train Precision: 0.9938
Test Precision: 0.8622
Train Accuracy: 0.9930
Test Accuracy: 0.8877
Train F1 Score: 0.9893
Test F1 Score: 0.8228

Random Forest Classifier with All Features and Tuned Hyperparameters
AUC: 0.8807
Train Recall: 0.5906
Test Recall: 0.5836
Train Precision: 0.8299
Test Precision: 0.8418
Train Accuracy: 0.8269
Test Accuracy: 0.8257
Train F1 Score: 0.6901
Test F1 Score: 0.6893

Random Forest Classifier with Feature Selection and Tuned Hyperparameters
AUC: 0.8850
Train Recall: 0.6340
Test Recall: 0.6272
Train Precision: 0.8097
Test Precision: 0.8203
Train Accuracy: 0.8320
Test Accuracy: 0.8309
Train F1 Score: 0.7112
Test F1 Score: 0.7108

AUC: 0.8850
Train Recall: 0.6340
Test Recall: 0.6272
Train Precision: 0.8097
Test Precision: 0.8203
Train Accuracy: 0.8320
Test Accuracy: 0.8309
Train F1 Score: 0.7112
Test F1 Score: 0.7108

8. Random Forest Classifier with feature selection and tuned hyperparameters.



CONCLUSION

NEXT STEPS

- Examine models for repeat guests
- Gather more data
- Contingency plan for false positives
- Financial impact report
- Speak with floor staff
- Surveys for guest cancellations
- Implement model in real time

QUESTIONS?

Norman Jen
normcjen3@gmail.com
https://www.linkedin.com/in/normanjen/

Yamuna Umapathy
u.yamuna@gmail.com
https://www.linkedin.com/in/yamuna-umapathy-b3a9081b2/