HOTEL BOOKING Classification

Presented by:

- Ajwad Almajnuni
- Rehab AL Zaidi
- Reem Alghamdi

Introduction

The purpose of this project is to predict if someone canceled the reservation or not by using Classification algorithms. We worked with data (HOTEL BOOKING) .

About Data:
This dataset contains observations for a City Hotel and a Resort Hotel. Each observation represents a hotel booking between the 1st of July 2015 and 31st of August 2017, including booking that effectively arrived.

It's contained:

- 32 Features .
- 119390 Observations .

Preprocessing

1

Check null values,
Duplicates,
Change Type

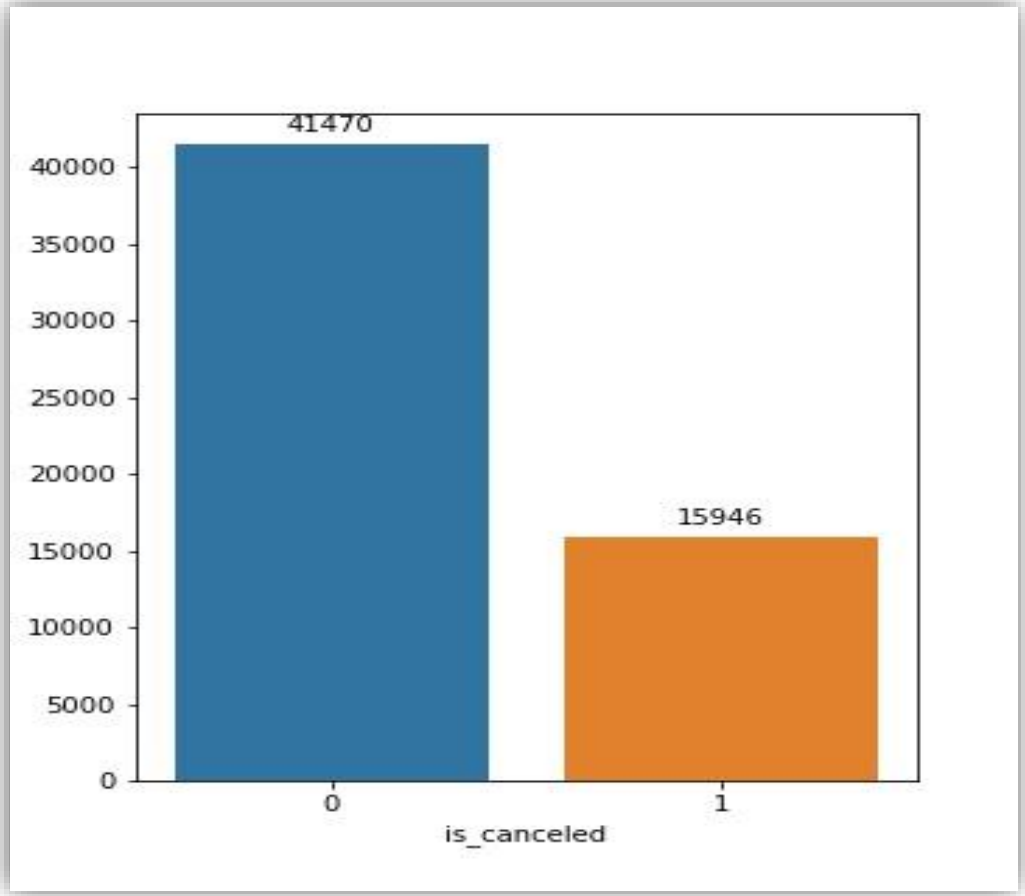
2

Get Dummies

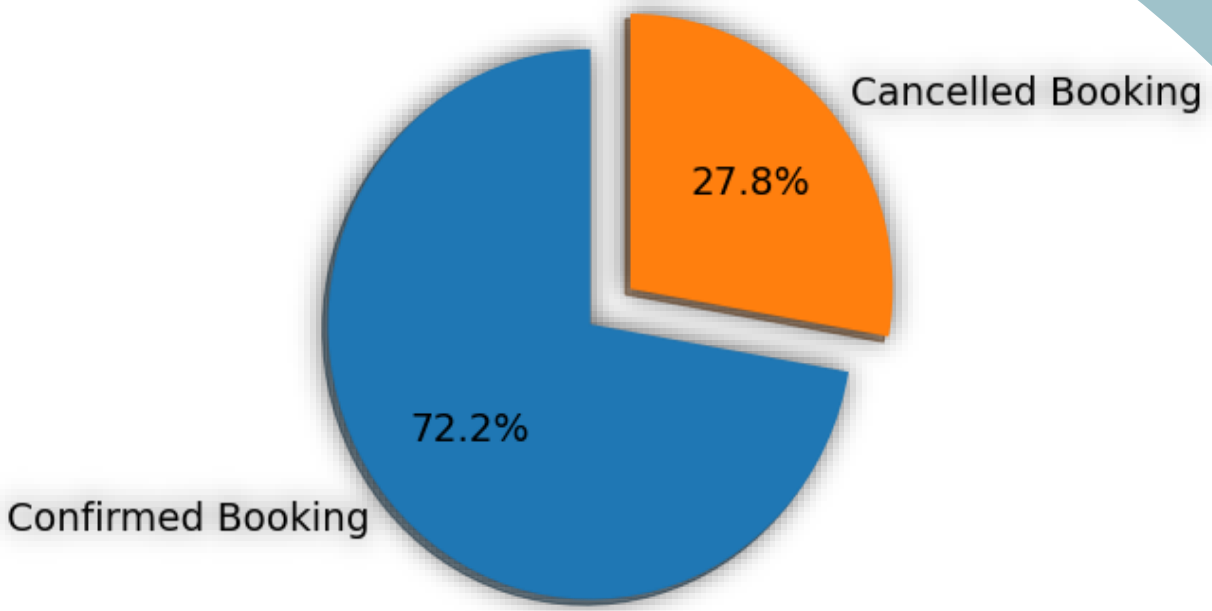
3

visualization

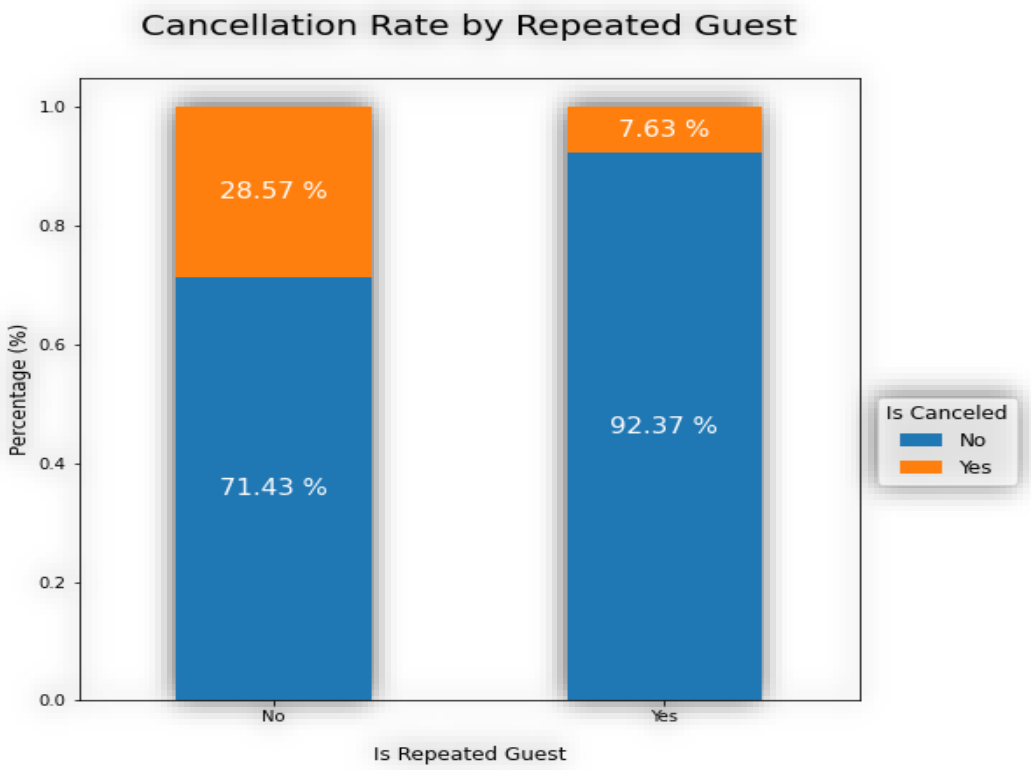
Classes imbalance



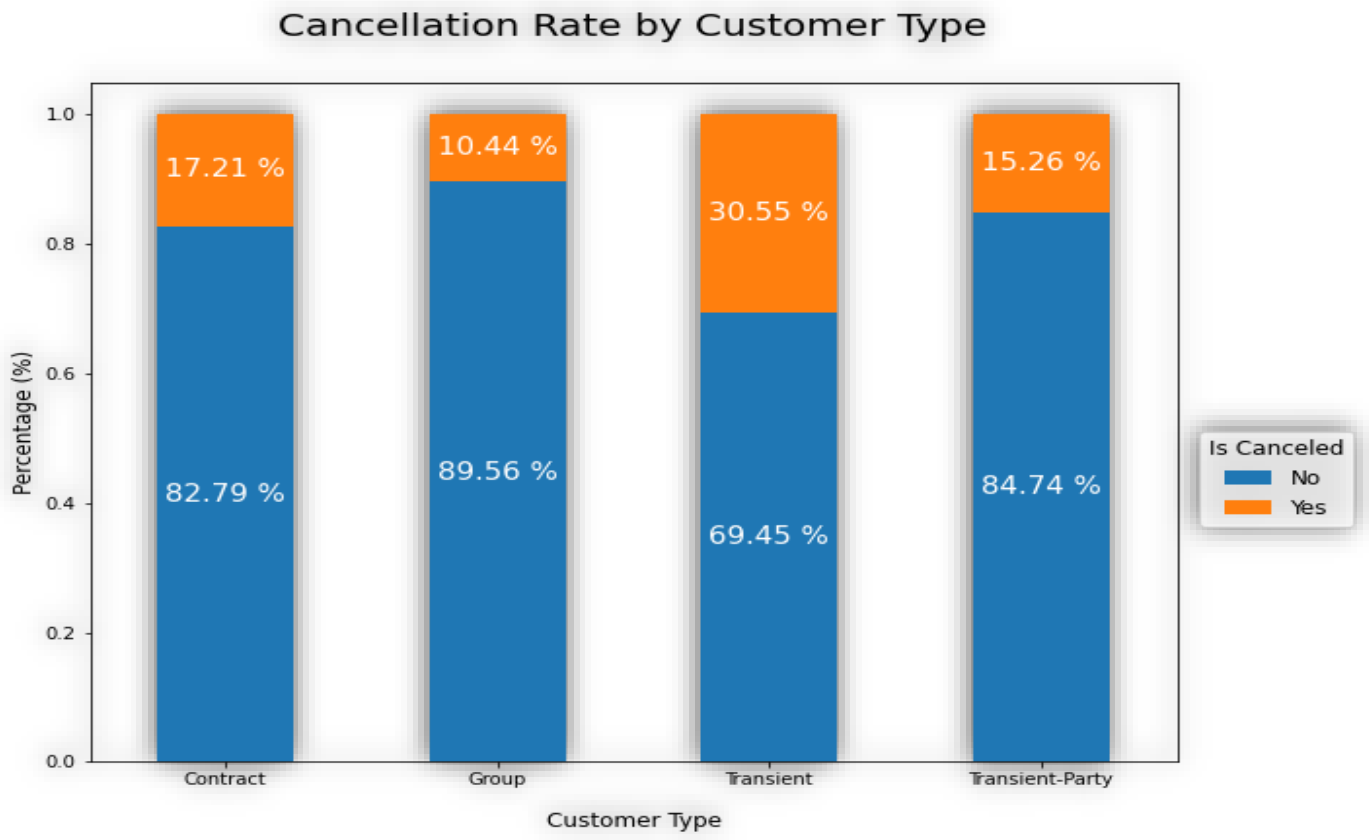
Cancellation Rate



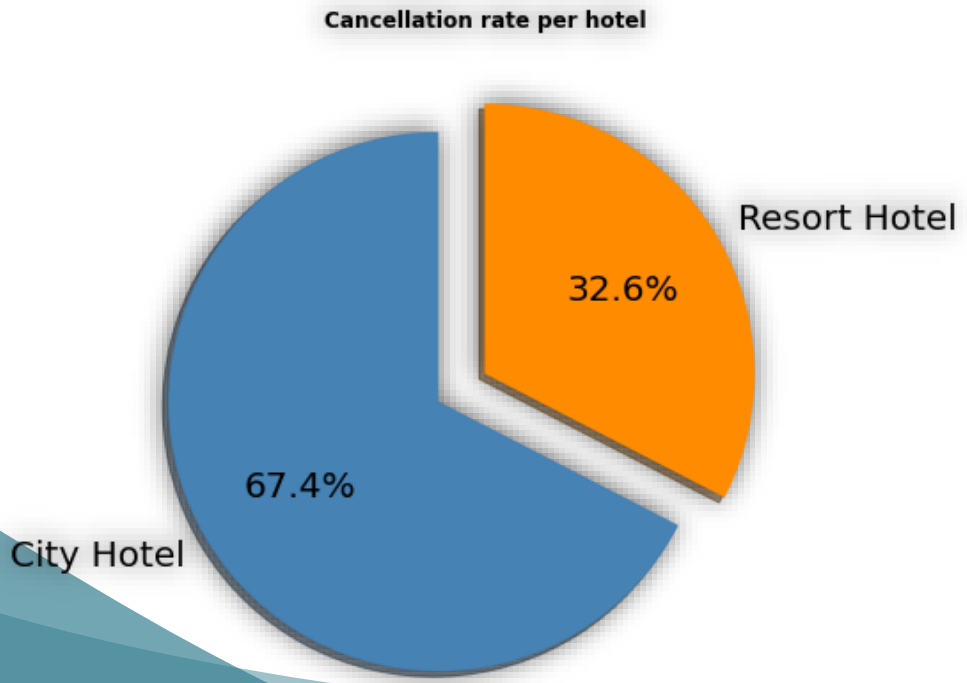
Repeated Guest and Cancellation



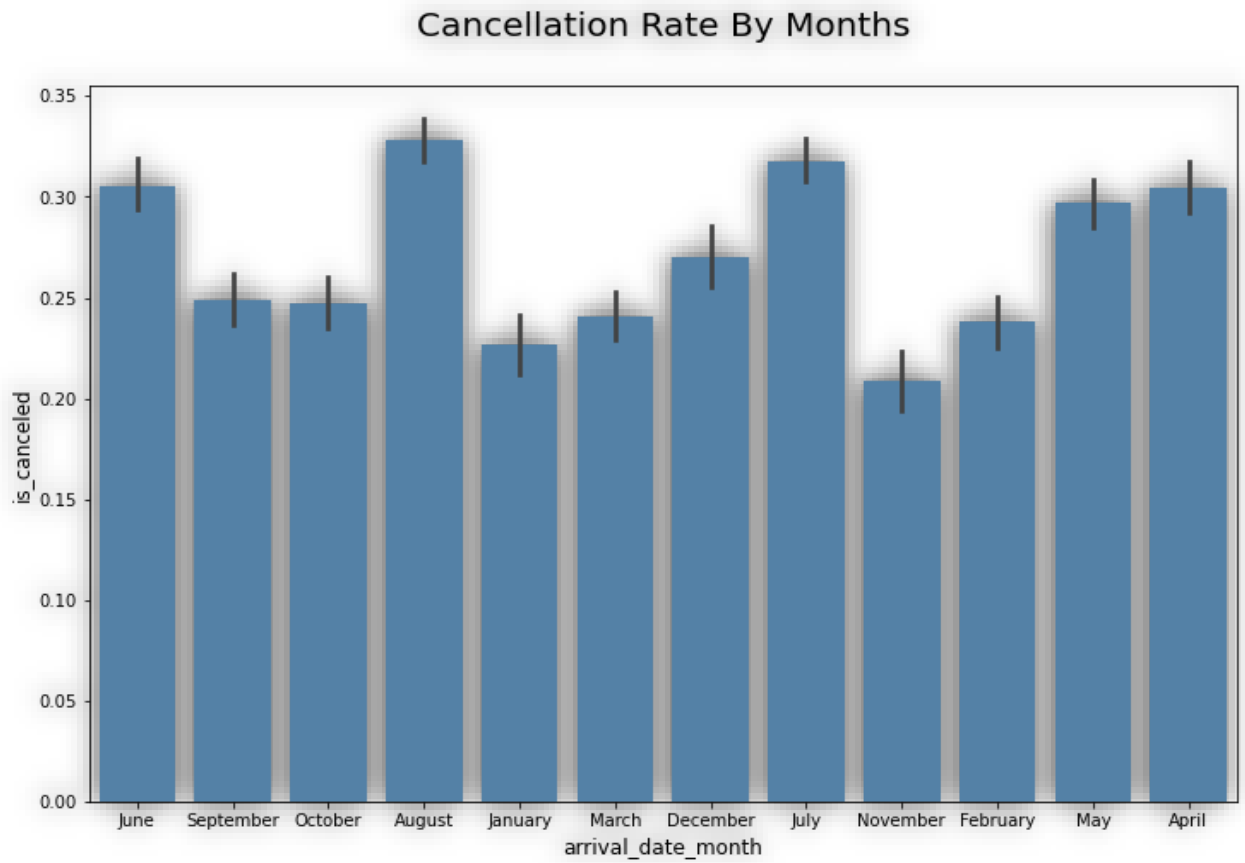
Cancellation Rate by Customer Type



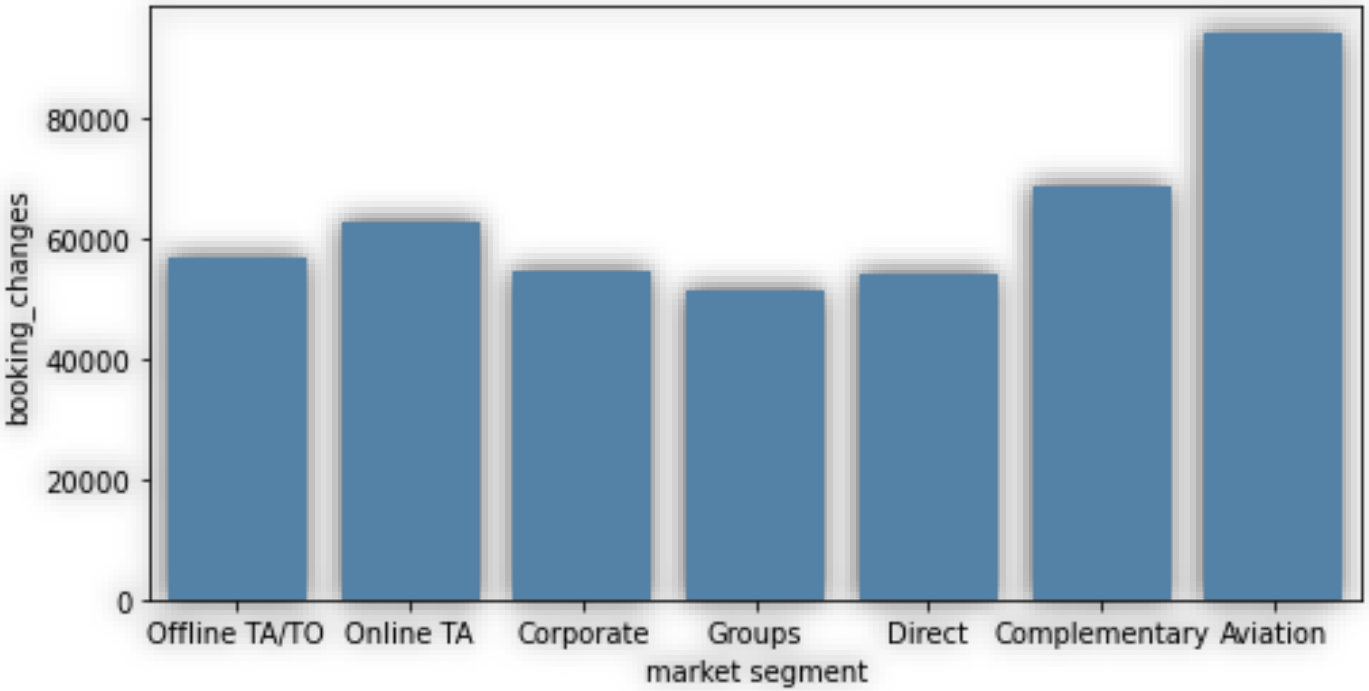
What is the cancellation rate for each hotel?



Cancellation Rate By Months



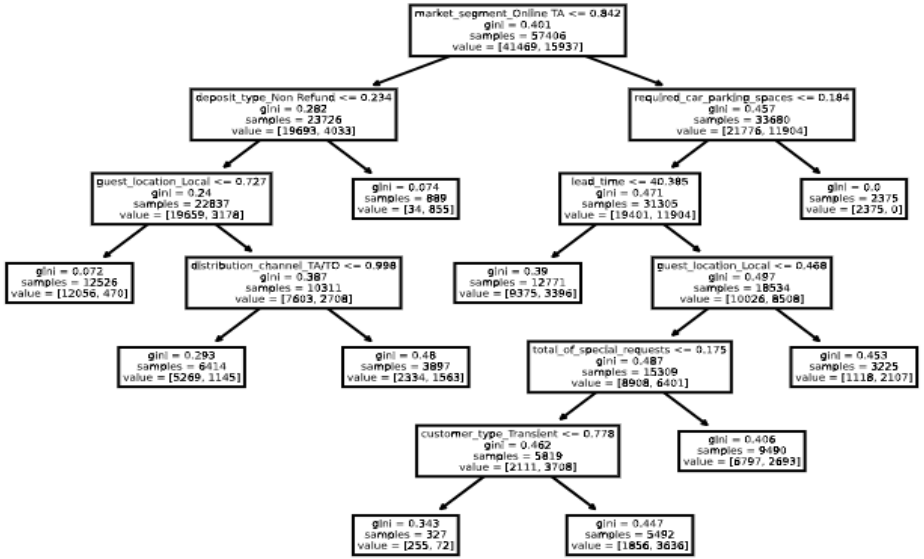
Is the number of changes of the reservation affected by (market segment)?



Models

Models	f1_score		
	Train	validation	Test
KNN	0.99	0.57	0.53
Logistic Regression	0.58	0.61	0.61
Random Forest	0.98	0.75	0.74
Decision Tree	0.09	0.21	0.18

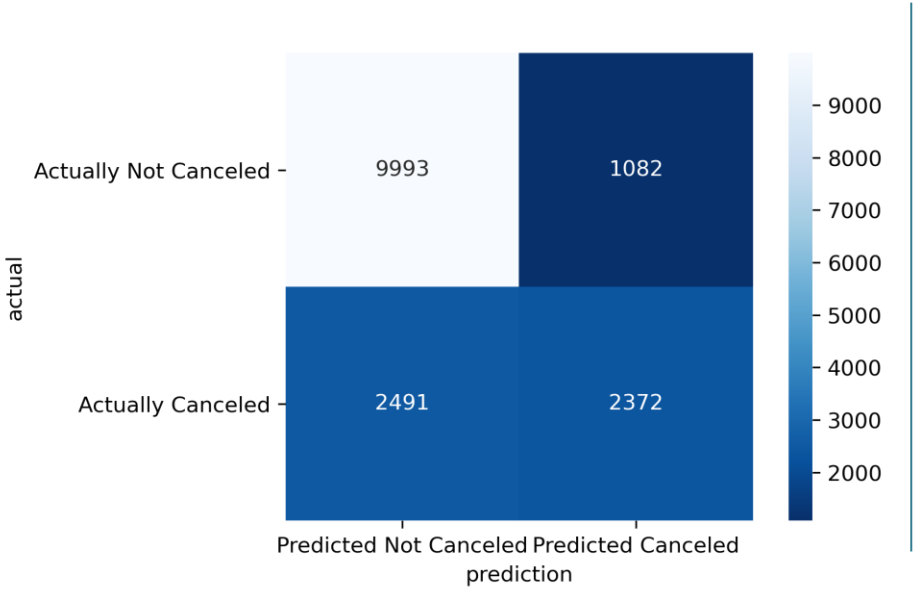
Decision Tree



Models

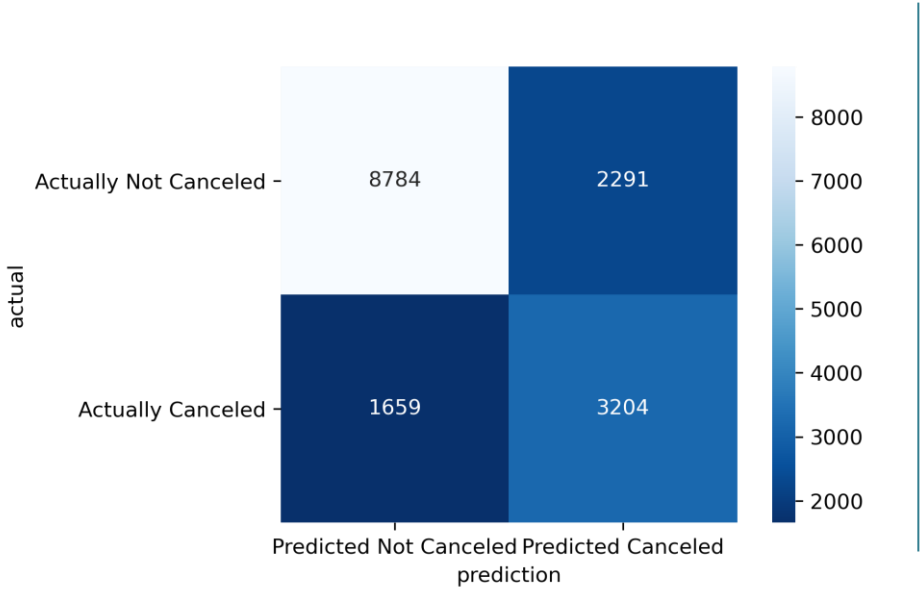
Models	f1_score		
	Train	validation	Test
Voting	0.99	0.74	0.73
Stacking Classifier	0.99	0.57	0.53
AdaBoost	0.35	0.43	0.42
XGBoost (Extreme Gradient Boosting)	0.75	0.74	0.73

KNN Evaluation



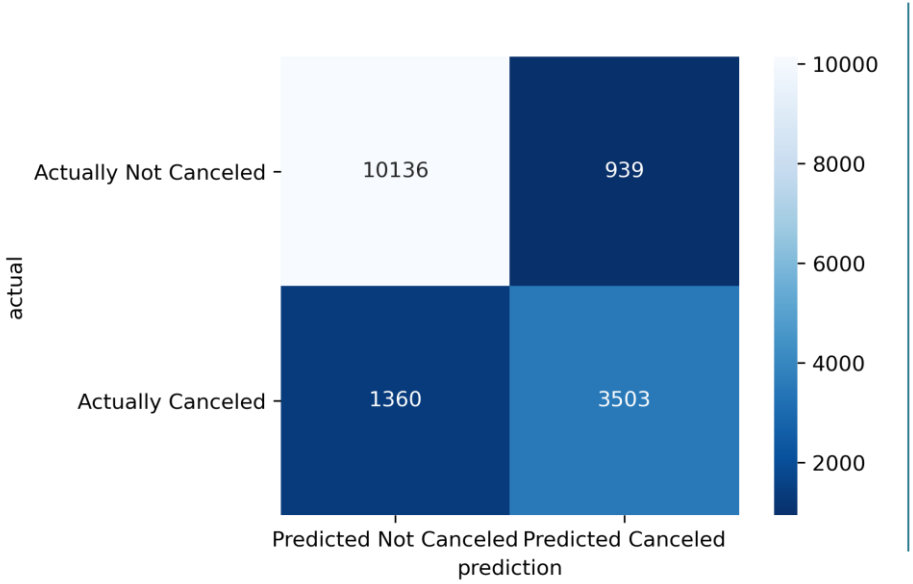
precision	recall
0.68	0.48

Logistic Evaluation



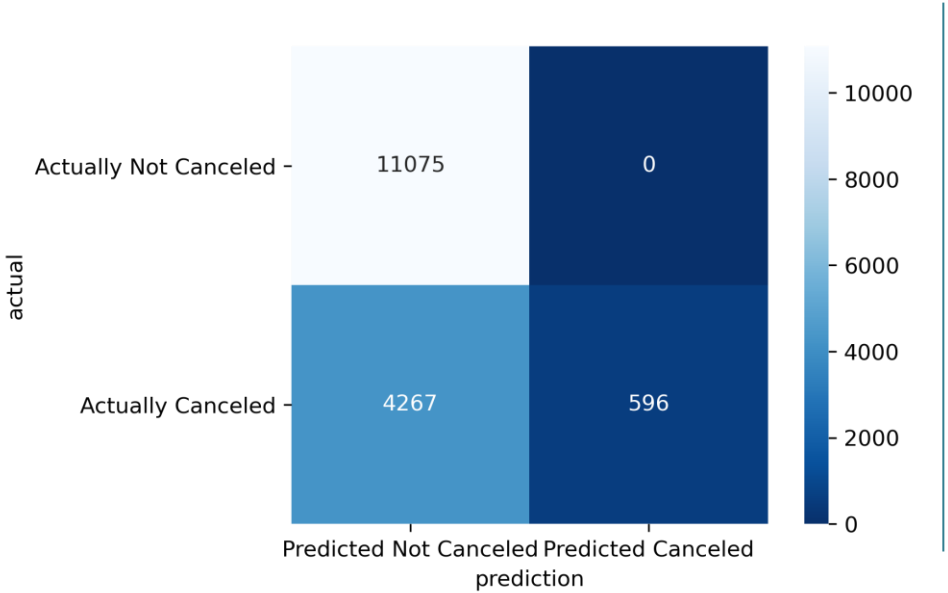
precision	recall
0.58	0.65

Random Forest Evaluation



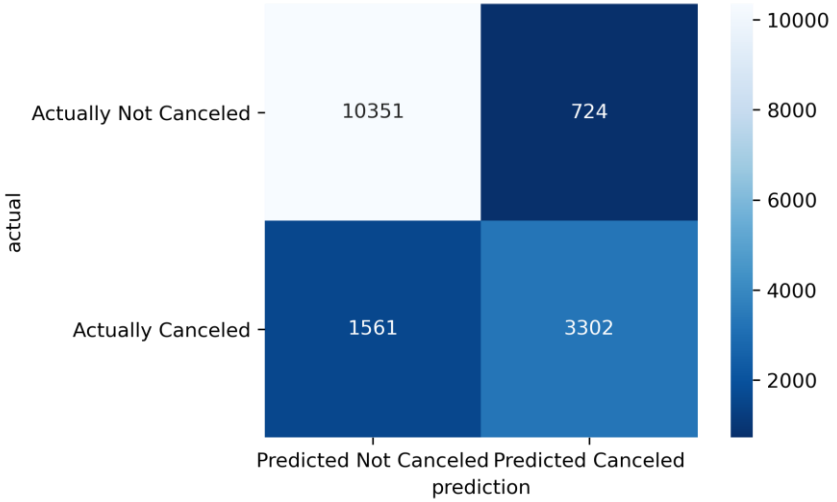
precision	recall
0.78	0.72

Decision Tree Evaluation



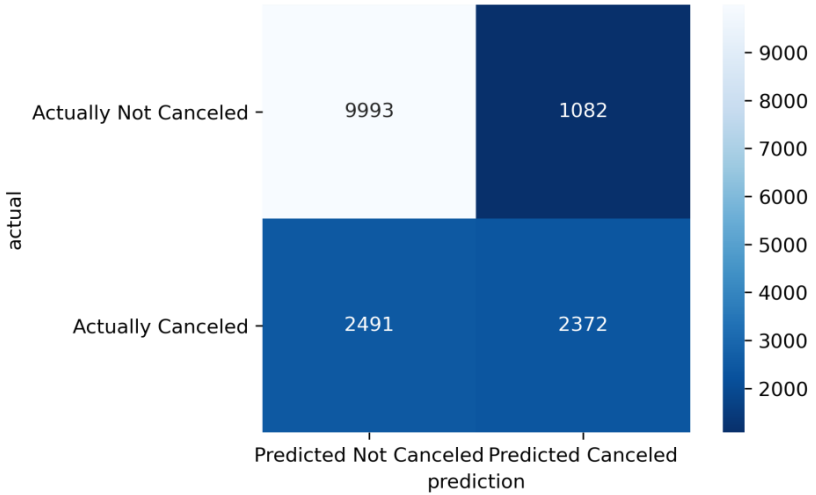
precision	recall
1.0	0.12

Voting Evaluation



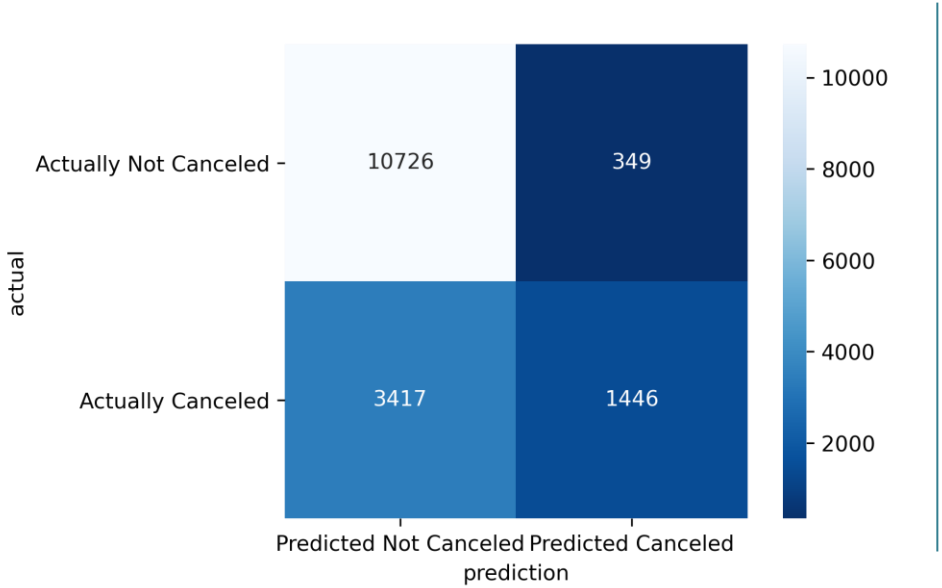
precision	recall
0.82	0.67

Stacking Evaluation



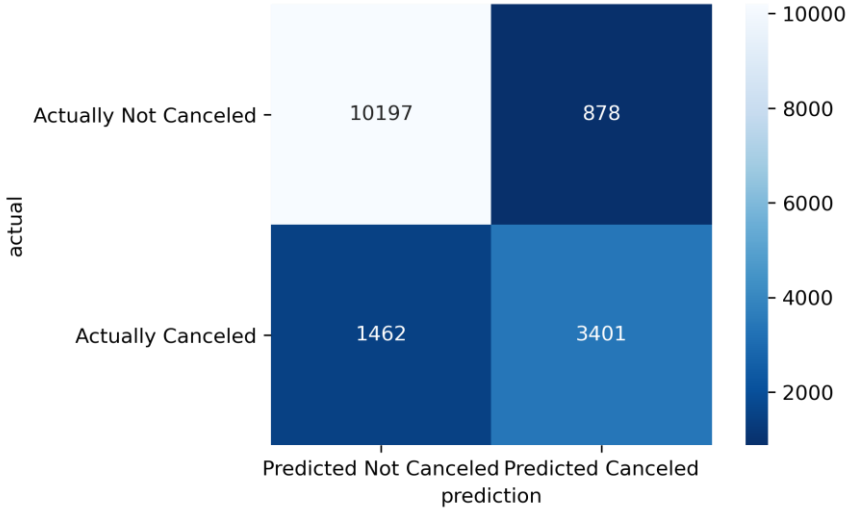
precision	recall
0.68	0.48

AdaBoost Evaluation



precision	recall
0.80	0.29

XGBoost (Extreme Gradient Boosting) Evaluation



precision	recall
0.79	0.69

Conclusion

In attempts to predict the best model if a person cancels the reservation or not, we made several models to determine the best model.

The best result is XGBoost (Extreme Gradient Boosting)

Train : **0.75**
Validation : **0.74**
Test : **0.75**

Thank You