Capstone

Hotel Booking Prediction

Presented by: Sharonda Petttiett - Warner

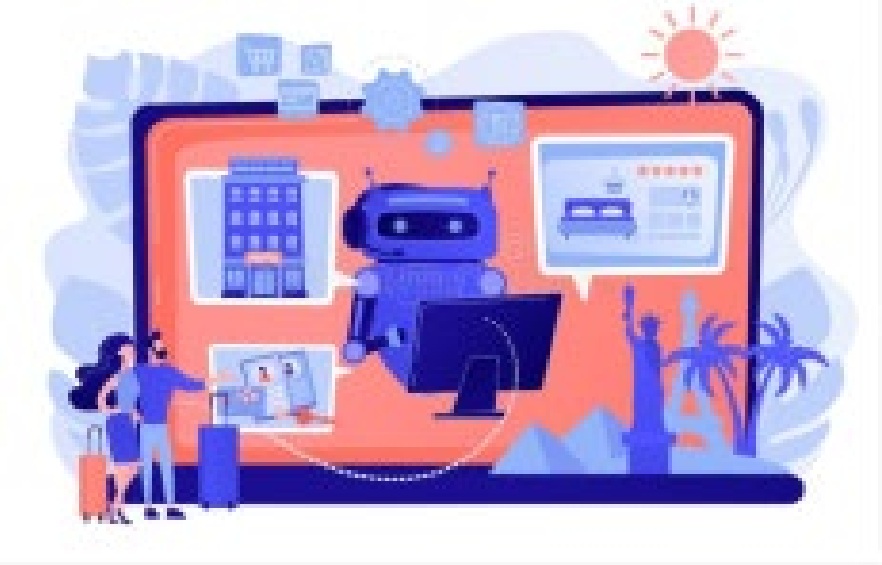
April 24, 2020

Problem Statement

- Reservations for resorts and hotels in Lisbon, Portugal are experiencing **37%** cancellations.
- Booking cancellations in the hospitality industry can result in:
  - Overbooking situations
  - Hotel's online social reputation
  - Revenue loss
  - Pricing, inventory and labor allocation decisions

Business Value

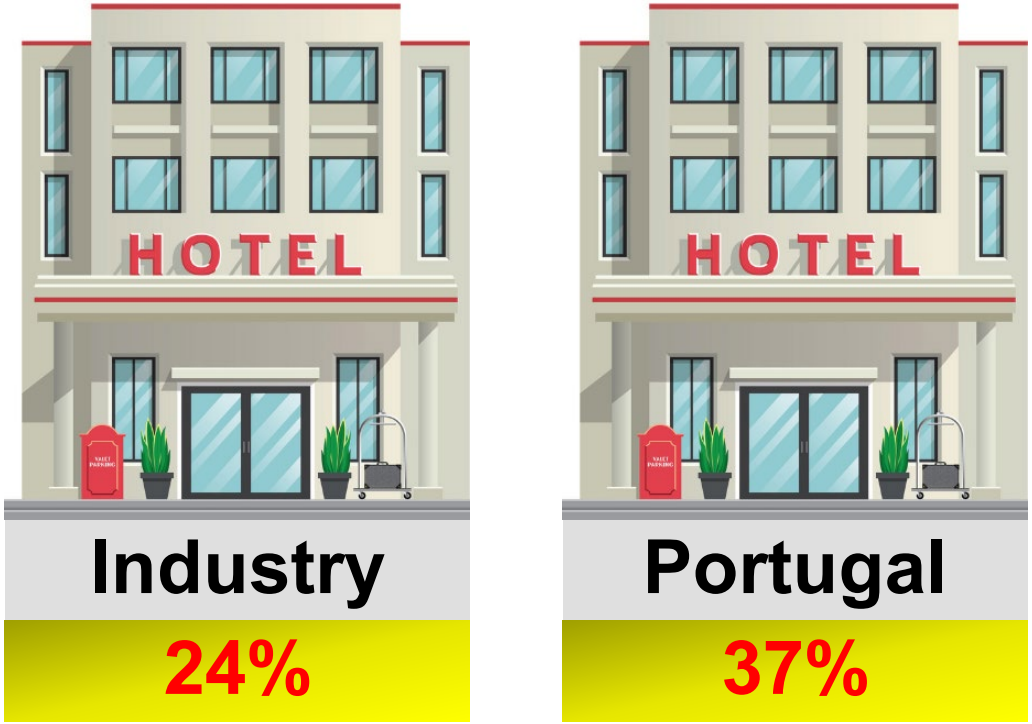
- To predict hotel booking cancellations to decrease uncertainty and increase revenue.
- To explain how future cancelled reservations can be predicted in advance by machine learning methods.



Methodology

- Classification problem, based on a Kaggle kernel
- Data Science concepts used:
  - Machine Learning (ML)
  - Hypothesis Testing
  - Data Visualization
  - Time Series

Industry Cancellation Rates



- Cancellation rate is **13%** higher than the Industry in Portugal

Loss Revenue

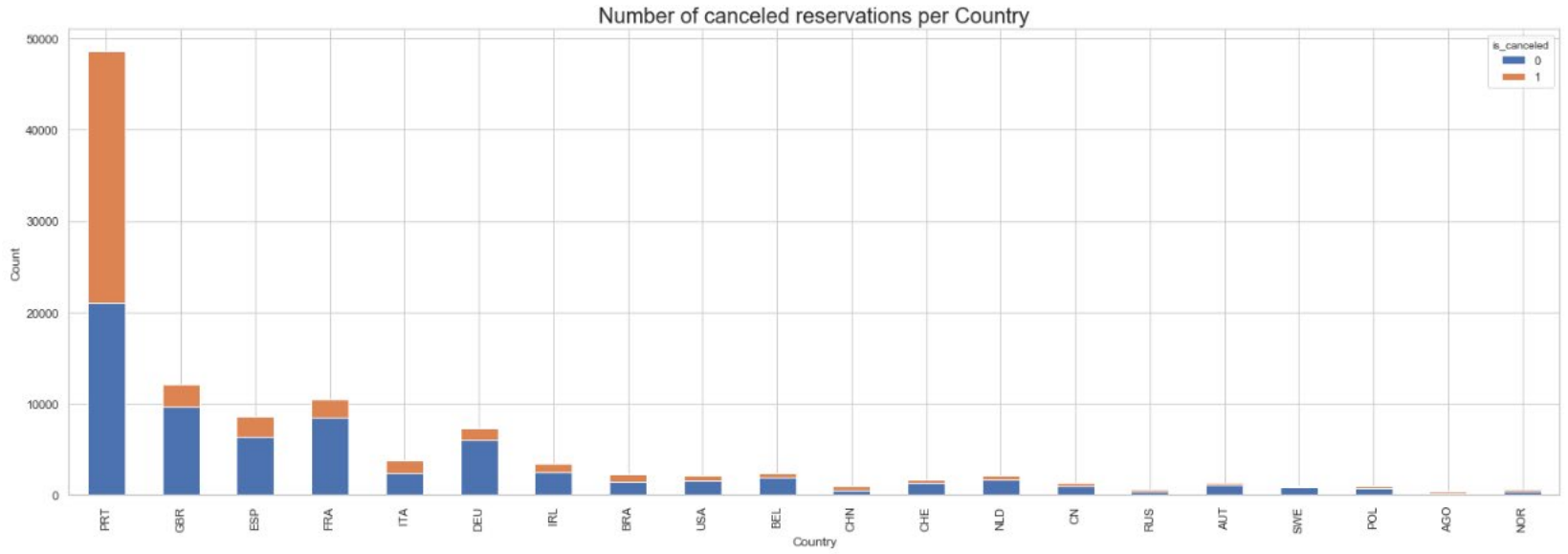
- Average lost from cancellations per year

Country

- Most of the cancellations are occurring in Portugal with 27,519

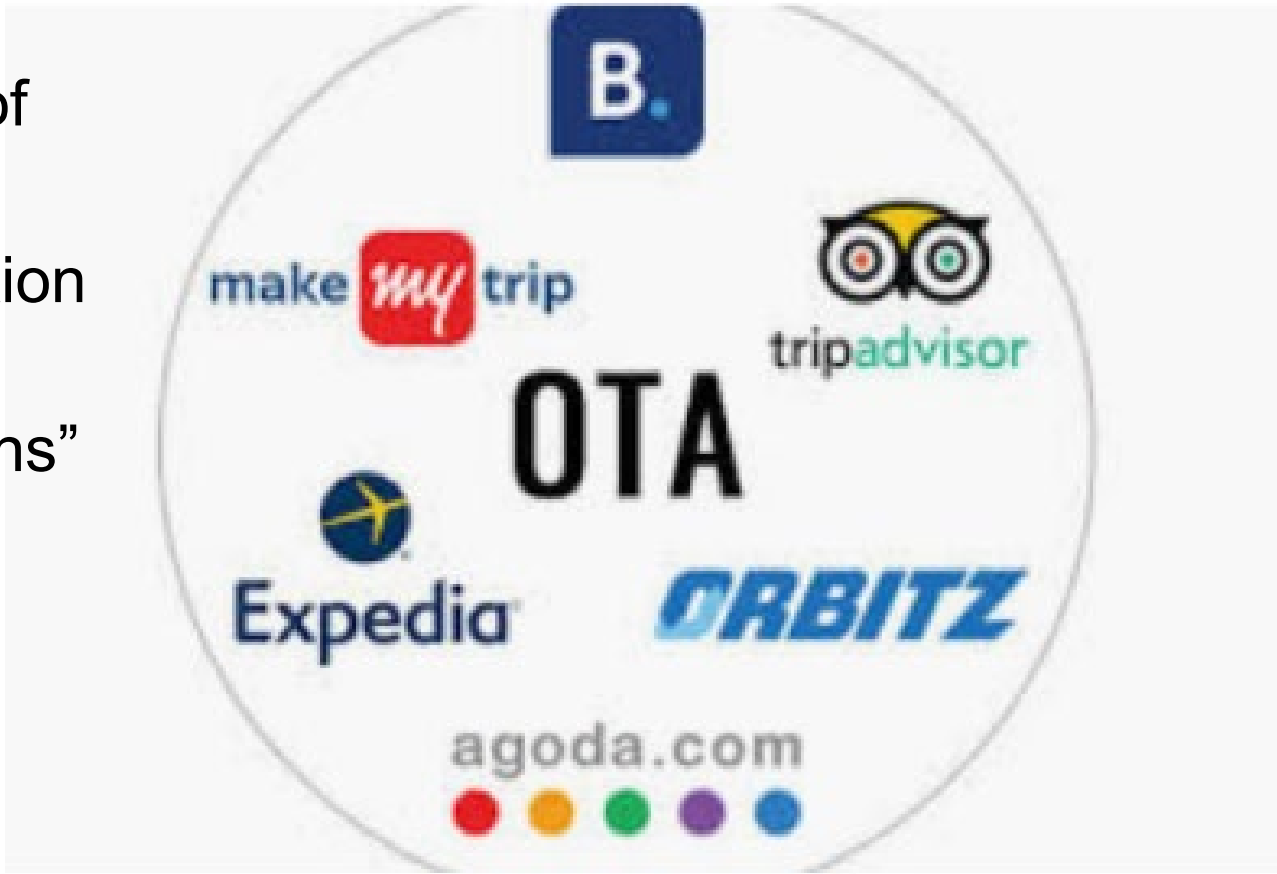
10x more than the following:

  - GBR 2,453
  - ESP 2,177
  - FRA 1,934

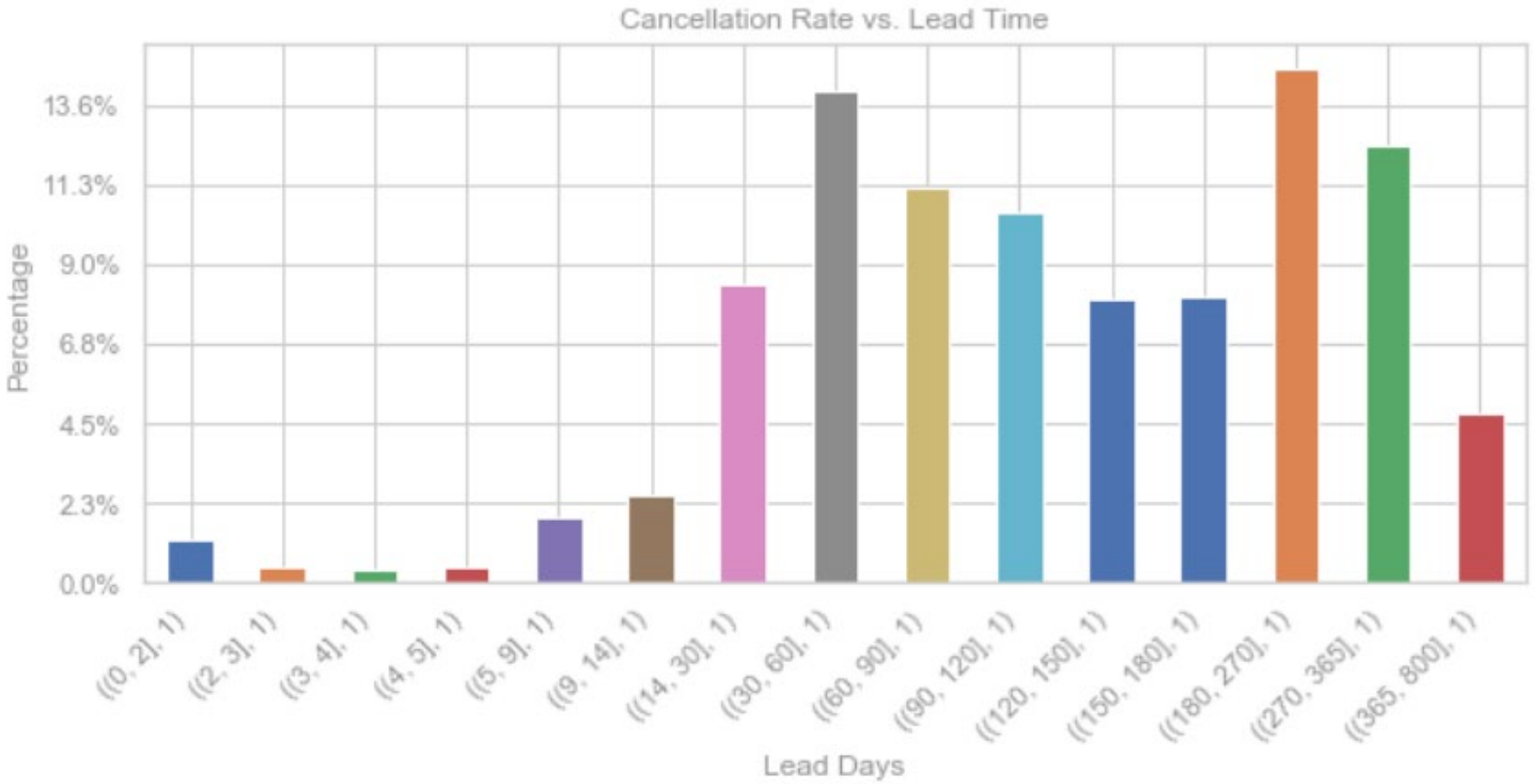


Online Tourist Agents (OTA)

- Make up **47%** of cancellations
- Industry cancellation rate is 40%
- “Free Cancellations” policies



Long Lead Times



- Longer than 30 days are **85%** more likely to be canceled

Findings - Most Predictive Features

- Number #1 predictor was '**Lead Time**'
- Others:
  - 'country',
  - 'market segment',
  - 'price',
  - 'customer type',
  - 'specific dates of reservation'

Findings – Best Model for Predictions

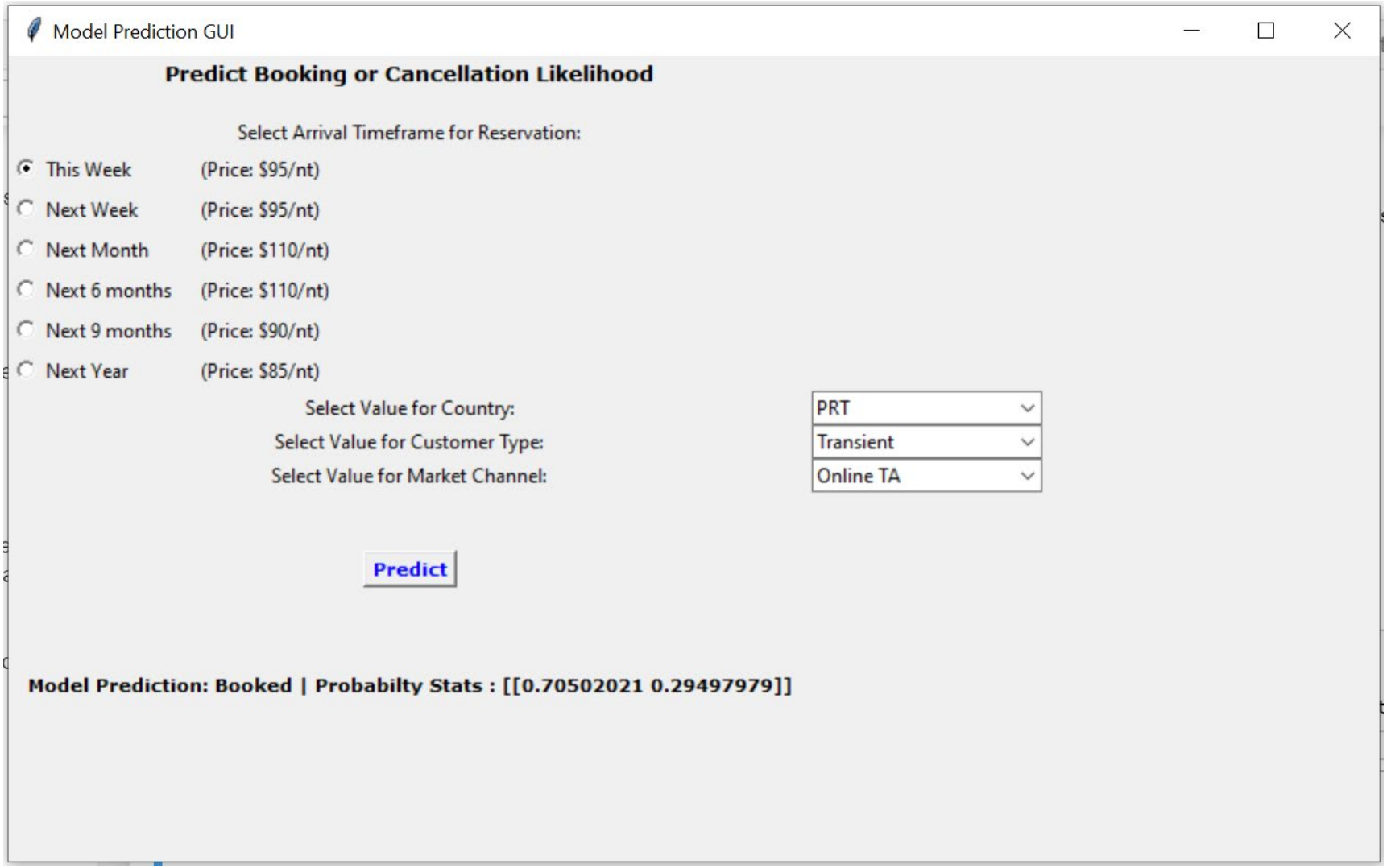
- Random Forest Classification have the "highest" scoring metrics

SCORING METRIC	VALUE
Test Accuracy Score	**88%**
Test Error Rate	**12%**
Precision	**87%**
ROC/AUC	**87%**
Sensitivity (Recall)	**81%**
Specificity	**92%**

Recommendations

- Hotels should analyze the growth rate of their respective marketing channels, esp.
  - OTA vs Direct
- Allow hotel managers to act on bookings with high cancellation probability
- Improve overbooking and cancellation policies

Let’s Do Some Predictions ….



Future Work

- Perform analysis between the city hotel and resort separately.
- Build a neural network to increase accuracy from 88% to 90%.
- Create a GUI to convert low cancellation probabilities into revenue.
- Perform analysis to determine which OTA will serves hotel better.
- Analyze data from weather, holidays and online prices/inventory.

Thank you!

Learn-co.slack.com

@ Sharonda W

LinkedIn: https://www.linkedin.com/in/sharonda-pettiett-8975693b/

Github: https://github.com/spettiett/dsc-capstone-project-v2-online-ds-pt-061019