INN Hotels Project

Context

A significant number of hotel bookings are called-off due to cancellations or no-shows. The typical reasons for cancellations include change of plans, scheduling conflicts, etc. This is often made easier by the option to do so free of charge or preferably at a low cost which is beneficial to hotel guests but it is a less desirable and possibly revenue-diminishing factor for hotels to deal with. Such losses are particularly high on last-minute cancellations.

The new technologies involving online booking channels have dramatically changed customers’ booking possibilities and behavior. This adds a further dimension to the challenge of how hotels handle cancellations, which are no longer limited to traditional booking and guest characteristics.

The cancellation of bookings impact a hotel on various fronts:

Loss of resources (revenue) when the hotel cannot resell the room.

Additional costs of distribution channels by increasing commissions or paying for publicity to help sell these rooms.

Lowering prices last minute, so the hotel can resell a room, resulting in reducing the profit margin.

Human resources to make arrangements for the guests.

Objective

The increasing number of cancellations calls for a Machine Learning based solution that can help in predicting which booking is likely to be canceled. INN Hotels Group has a chain of hotels in Portugal, they are facing problems with the high number of booking cancellations and have reached out to your firm for data-driven solutions. You as a data scientist have to analyze the data provided to find which factors have a high influence on booking cancellations, build a predictive model that can predict which booking is going to be canceled in advance, and help in formulating profitable policies for cancellations and refunds.

Data Description

The data contains the different attributes of customers' booking details. The detailed data dictionary is given below.

Data Dictionary

Booking\_ID: unique identifier of each booking

no\_of\_adults: Number of adults

no\_of\_children: Number of Children

no\_of\_weekend\_nights: Number of weekend nights (Saturday or Sunday) the guest stayed or booked to stay at the hotel

no\_of\_week\_nights: Number of week nights (Monday to Friday) the guest stayed or booked to stay at the hotel

type\_of\_meal\_plan: Type of meal plan booked by the customer:

Not Selected – No meal plan selected

Meal Plan 1 – Breakfast

Meal Plan 2 – Half board (breakfast and one other meal)

Meal Plan 3 – Full board (breakfast, lunch, and dinner)

required\_car\_parking\_space: Does the customer require a car parking space? (0 - No, 1- Yes)

room\_type\_reserved: Type of room reserved by the customer. The values are ciphered (encoded) by INN Hotels.

lead\_time: Number of days between the date of booking and the arrival date

arrival\_year: Year of arrival date

arrival\_month: Month of arrival date

arrival\_date: Date of the month

market\_segment\_type: Market segment designation.

repeated\_guest: Is the customer a repeated guest? (0 - No, 1- Yes)

no\_of\_previous\_cancellations: Number of previous bookings that were canceled by the customer prior to the current booking

no\_of\_previous\_bookings\_not\_canceled: Number of previous bookings not canceled by the customer prior to the current booking

avg\_price\_per\_room: Average price per day of the reservation; prices of the rooms are dynamic. (in euros)

no\_of\_special\_requests: Total number of special requests made by the customer (e.g. high floor, view from the room, etc)

booking\_status: Flag indicating if the booking was canceled or not.