Reservation Cancelation Dataset

Competition Link

https://www.kaggle.com/competitions/playground-series-s3e7

Dataset Description

The dataset for this competition (both train and test) was generated from a deep learning model trained on the Reservation Cancellation Prediction dataset. Feature distributions are close to, but not exactly the same, as the original. Feel free to use the original dataset as part of this competition, both to explore differences as well as to see whether incorporating the original in training improves model performance.

Files

- **train.csv** - the training dataset; `booking_status` is the target (e.g., whether the reservation was cancelled)
- **test.csv** - the test dataset; your objective is to predict `booking_status`
- **sample_submission.csv** - a sample submission file in the correct format

About Dataset

Customer behaviour and booking possibilities have been radically changed by online hotel reservation channels. Cancellations or no-shows cause a significant number of hotel reservations to be cancelled. Cancellations can be caused by a variety of factors, such as scheduling conflicts, changes in plans, etc. In many cases, this is made easier by the possibility of doing so free or at a low cost, which is beneficial for hotel guests but less desirable and possibly revenue-diminishing for hotels.
As a Data Scientist, your job is to build a Machine Learning model to help the Hotel Owners better understand if the customer is going to honour the reservation or cancel it ?

Dataset Description

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

- Booking_ID: unique identifier of each booking
- No of adults: Number of adults

- No of children: Number of Children
- noofweekend_nights: Number of weekend nights (Saturday or Sunday) the guest stayed or booked to stay at the hotel
- noofweek_nights: Number of week nights (Monday to Friday) the guest stayed or booked to stay at the hotel
- typeofmeal_plan: Type of meal plan booked by the customer:
- requiredcarparking_space: Does the customer require a car parking space? (0 - No, 1- Yes)
- roomtypereserved: 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)
- noofprevious_cancellations: Number of previous bookings that were canceled by the customer prior to the current booking
- noofpreviousbookingsnot_canceled: Number of previous bookings not canceled by the customer prior to the current booking
- avgpriceper_room: Average price per day of the reservation; prices of the rooms are dynamic. (in euros)
- noofspecial_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.