Neural Network
Project-3

HOTEL BOOKING PREDICTION USING DEEP LEARNING

An Online travel booking company is suffering from loss in revenue because of the uncertain booking cancelation of its customers. The company wants to know which customer will cancel the booking. As a data-scientist we have to help the company to predict whether the customer will cancel the booking or not. We have all the booking details like arrival_date_year, stays_in_week_nights, arrival_date_day_of_month etc of the customers from various countries. We have to do some data analysis to answer some questions and we have to run a NN model to predict whether the customer will cancel the booking or not.

DataSet Description:

hotel	Type of hotel(resort hotel,city hotel)
is_cancel (Target)	Is the booking is cancel or not
arrival_date_year	Year of arrival
Arrival_date_month	month the guest arrives
Arrival_date_day_of _month	date the guest arrives
Stays_in_weekend_nights	Weekend night guest spends in hotel
Stays_in_week_nights	Weekdays night guest spends in hotel
adults	No.of adults
children	No. of children
babies	How many babies the guest have
meals	Type of meal
country	country in which the hotel located (Example PRT-Portugal, IDN-India,GBR- Great Britain)
distribution_channel	Distribution type of guest

is_repeated_guest	Whether the guest previously stayed in the same hotel or not

previous_cancelation	Whether the guest previously cancel the booking
previous_booking_not_canceled	Whether the guest previously not cancel the booking
reserved_room_type	Type of room the guest reserved
assigned_room_type	Type of room that is assigned to the guest
deposit type	Deposit type of guest
days_in_waiting_list	Waiting days for the guest
customer_type	What type of guest
required_car_parking_spaces	How much car parking space required
reservation_status	Whether the guest has checkout or cancel

EXPLORATORY DATA ANALYSIS(EDA)(15 Points)

Q1. Read the dataset and visualize the target(i.e. is_cancel). State whether it is imbalanced or not. How we can deal with class imbalance, state briefly.(3 point)

Split the dataset into 80:20 ratio.(80% training and 20% testing).Visualize the test

data. NOTE: ALL QUESTIONS FROM Q2. to Q6. will be answered on the test data.

Q2. In the test data in which type of hotel the cancelation is more?(2 points)

Q3. Which Country has more resort hotels and which country has more city hotels?(3 points)

Q4. How many check-outs has been done in hotels in India, Here IDN, refers to India?(2 Points)

Q5. In which countries the minimum number of BB & SC meals have been booked?(3points)

Q6. State the inferences between deposit_type and is_cancel.(2 points)

Q7. Please follow below steps

7.1 Experiment with different neural networks with parameters combinations given in the below table and run each model. (4 points)

7.2 State the difference between results obtained in train and test dataset clearly with reasons. (2 points)

7.3 Please drop the "reservation status" column and follow the below steps. (1 point)

7.4 Experiment with different neural networks with parameters combinations given in the below below and run each model. (**Note:- the dataset is without "reservation status"**) (4 points)

7.5 State the difference between results obtained in train and test dataset (**"excluding reservation status"**) clearly with reasons. State which model we should use for production. (4 points)

No. of Hidden Nodes	Learning Rate	Momentum	Normalizer
128	0.0001	0.3	Do not Normalize
64	0.001	0.2	Mini-Max Normalizer
32	0.03	0.5	Gaussian Normalizer