Predicting the cancellation of hotel booking?

January 21, 2023

1 Description

The online hotel reservation channels have dramatically changed booking possibilities and customers' behavior. A significant number of hotel reservations are called-off due to cancellations or no-shows. The typical reasons for cancellations include change of plans, scheduling conflicts, etc [1]. 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 [2]. Can you predict if the customer is going to honor the reservation or cancel it?

Figure 1: Let's predict it into two classes (Not Canceled vs Canceled).

Attribute Information:
Dataset consists of a CSV which have a record of 36,275 booking details with their 18 corresponding features and a target variable (i.e. Booking Status) indicating if the booking status was Not Canceled vs Canceled. The classes to be predicted as given as follows:

1. Not Canceled
2. Canceled

2 Material

Your dataset is available to be downloaded Click me.

3 Aim

Your job is to create a binary classification model to classify hotel booking into (i.e. Not Canceled vs Canceled)

4 Help

- Try different classification algorithms such as SVM, Random Forest, KNN, and more [3]
- Split for data into training and validation. You have a separate test set to evaluate your model.
- Feel free to take your project in any direction. However, make sure you can justify that the approach is leading to better prediction performance.

- Exploratory data analysis is quite important and please use pandas.
- Please choose the correct evaluation metrics.
- In case of any doubt, please discuss with CE880 academic staff (Dr Haider Raza).

References

[1] N. Antonio, A. de Almeida, and L. Nunes, "Predicting hotel bookings cancellation with a machine learning classification model," in *2017 16th IEEE International Conference on Machine Learning and Applications (ICMLA)*, 2017, pp. 1049–1054.

[2] A. J. Sánchez-Medina, C. Eleazar *et al.*, "Using machine learning and big data for efficient forecasting of hotel booking cancellations," *International Journal of Hospitality Management*, vol. 89, p. 102546, 2020.

[3] C. Robert, "Machine learning, a probabilistic perspective," 2014.