Predictive Model for Hotel Booking Demand

Fagun Shah-fagun.shah@ryerson.ca

Abstract

The hotel industry has been successfully operating for centuries now. For the development of tourism opportunities hospitality plays a vital role. Also, in order to provide the guests with best services and amenities it is mandatory for the hotels to predict the flow of the passengers. The data can be found in the article "Hotel booking demand datasets" written by Nuno Antonio, Ana Almeida, and Luis Nunes for Data in Brief, Volume 22, February 2019. The data was later downloaded and cleaned by Thomas Mock and Antoine Bichat for TidyTuesday an online weekly event during the week of February 11th, 2020, I shall be using this cleaned and labelled dataset for analysis. Present data is extracted from the SQL databases of a city hotel and a resort hotel in Portugal with variables like arrival and departure date, number of adults, children and babies, meal, country, deposit type, agent, company, etc. The main goal of the project is to build a predictive model to predict the hotel booking cancellations. The plan is to use various tree-based algorithms and pick the model having highest accuracy.

References:

1. N. Antonio, A. De Almeida, and L. Nunes, "Hotel booking demand datasets", Data in brief, vol. 22, 2019
2. T. Mock and A. Bichat, TidyTuesday master data available at : https://github.com/rfordatascience/tidytuesday/blob/master/data/2020/2020-02-11/readme.md