**Enhance the guest experience for Indian Hotel Industry using a predictive guest rating model**

*For Cousera Capstone Project*

Anurag Nirmal

March 07, 2020

1. **Introduction**
   1. **Background**

The Indian tourism and hospitality industry have emerged as one of the key drivers of growth among the services sector in India. Tourism in India has significant potential considering the rich cultural and historical heritage, variety in ecology, terrains and places of natural beauty spread across the country.

The Indian hotel industry is one of most digital advanced industry where a variety of digital tools are used for planning, booking and providing online feedback. The industry consists of international hotel chains, start-ups like Airbnb, Oyo, etc and fragmented small & medium hotels. It is important to understand the factors contributing to a guest (user or customer who stay at the hotel) experience. Efforts taken to enhance the experience will directly lead to higher business revenue, repeat business opportunity, network effect on other travellers and growth of the overall growth of the tourism industry.

* 1. **The Problem Statement**

The Indian hotel industry is big and dynamic comprising of big, medium and small players as per their star ratings. While the hotel star rating is important, the experience the guest or customer is visible in the hotel user rating and number of reviews left by the users. A good stay at the hotel is reflected in 5-star rating and positive review. A bad stay at the hotel is reflected in low start rating and negative review.

The project aims to understand the factors responsible (for example, hotel brand, location, point of interest, room area, etc) and their relative importance to determine the user rating. Also, comparison will be drawn against the rating provided by the website (Goibibo.com).

* 1. **Interest in the Project**

The customer for this project who will have the maximum interest are the following

a) Hotel owners may use the model to understand the feature set and their relative importance in predicting the hotel user rating. They can focus on specific factors to improve their hotel and enhance their revenue in the process

b) Hotel website may use the model to position the hotels on the webpage which have more model rating and hence likelihood to provide better experience to prospective customers

1. **Data Acquisition and Cleaning** 
   1. **Data Sources**

A pre-crawled dataset (of 4000 Indian hotels) which was created by extracting data from goibibo.com, a leading travel site from India. This dataset is available at **Kaggle.com.**

* 1. Data Cleaning

*To be covered in next week*

* 1. Feature Selection

*To be covered in next week*

1. Exploratory Data Analysis
   1. Calculation of Target Variable
   2. Inferences from the Data
2. Result of the Study
   1. The Model - XYZ
   2. Comparison of Different Models
   3. Recommendations
3. Conclusion
4. Future Study and Interest