Title:

Comparative analysis of machine learning models for Hotel Price Prediction

PROBLEM STATEMENT

The problem statement was derived from the existing knowledge of online hotel booking system, supervised machine learning techniques, and a web scraped dataset from ‘Airbnb price prediction’ dataset on Kaggle. The given dataset has a 74,000+ records and about 29 columns containing all the necessary details about hotels across various cities. We need to train a machine learning model in such a way that given the details of a hotel it can predict the price of the hotel from these details.

An accurate forecast of occupancy and room revenue empowers a revenue manager to yield across various channels. For example, if a property is forecasted for a high occupancy owing to high unconstrained demand, then the revenue manager can choose to yield and sell on low cost/high-rate channels to maximize profits.

This project will help anyone looking for an ideal hotel within their price limit just by specifying the details, and will thus save their time by providing all other relevant details that match with the costumer’s choice, in one go. In order to approach the problem, regression analysis will be used.

Exploratory data analysis formulation of a prediction model using machine learning algorithms and comparative analysis.

ABSTRACT

The project aims at forecasting the price of a hotel based on its historical data such as review rating score, where the hotel is located, amenities provided by the hotel, apartment type, and many more. Exploratory data analysis, supervised machine learning models such as Linear Regression, Logistic Regression and decision trees will be implemented in this project to visualize patterns and find a relevant relation between the historical data and values that are to be predicted. The models will be trained and implemented with boosting algorithms such as AdaBoost and XGBoost to further improve the accuracy of the trained models. These models finally be tested and compared for their accuracy. The most accurate model obtained shall conclude the project.