greatlearning
Learning for Life

# Project Summary

| | |
|---|---|
| Batch details | DSE CHN Apr'23 |
| Team members | Mohammed Anaz, D K Tarun Venkatesh, Malini, Vishal Raj, Sangeetha Elangovan |
| Domain of Project | Travel and Tourism |
| Proposed project title | Hotel Booking Cancellation Prediction |
| Group Number | 2 |
| Team Leader | Vishal Raj |
| Mentor Name | Ankush Bansal |

Date: 03.09.2023

Ankush Bansal

Signature of the Mentor

Vishal Raj

Signature of the Team Leader

greatlearning
Learning for Life

# Table of Contents

greatlearning
Learning for Life

# 1. OVERVIEW

The goal of this machine learning project is to develop a predictive model that can accurately predict whether customers who book hotel rooms will show up for their reservation or cancel it at the last minute. By achieving this objective, we aim to assist hotels in optimizing their operations, improving resource allocation, and reducing revenue losses due to cancellations.

**Expected Benefits:**

- Improved resource allocation for hotels, reducing operational costs.
- Enhanced customer experience by reducing booking cancellations.
- Increased revenue through optimized room management.
- Data-driven insights for hotel management to make informed decisions.

# 2. BUSINESS PROBLEM STATEMENT

## 1. Business Problem Understanding

The business problem underlying the hotel booking dataset is to gain insights into booking patterns and customer behavior for two types of hotels: city hotels and resort hotels.

## 2. Business Objective

Develop a Booking Cancellation prediction system that allows hotel owners to whether the guests who made the booking will actually come or not. Reduce the risk of knowing the cancellation at the last moment and be prepared with the prediction system.

## 3. Approach

By studying the previous data of customer bookings and building a model using machine learning techniques to predict the actual footfall of customers who are booked for the Hotel stay.

4. **Conclusions**

By implementing the Trained model, our aim is to create a predictive model with good accuracy to detect whether the booked customers are actually going to appear on the booked date.

# 3. TOPIC SURVEY IN BRIEF

1. **Problem understanding**

During the time of year when there is a radical number of bookings due to holidays and events this model will help us determine if the customer who booked the rooms will show up (pay for the rooms after stay) at these peak seasons or the reserved rooms would not bring in revenue.

The model is about analyzing a dataset that contains two distinct hotels namely "city hotels" and "resort hotels", The primary objective is to find whether the person who booked hotel rooms, checked in or cancelled the booking. The challenge is significant for hotel management as it helps in determining the status during the period of booking and helps to manage their resources effectively and enhance their operations.

2. **Current solution to the problem**

Currently, there might be existing systems or methods like manual tracking or basic software to track booking however these might lack the ability to track and predict things accurately.

- **Reservation System:**

Hotels have reservation systems in place that record all bookings made by guests. When a guest makes a reservation, their information is stored in the system. This information includes the guest's name, contact details, check-in/check-out dates, room type, and any special requests.

- **Cancellation Process:**

If a guest needs to cancel their booking, they typically do so through the hotel's reservation system or by contacting the hotel directly. The cancellation policy and any associated fees are applied according to the terms and conditions of the booking.

- **Communication:**

Hotels often communicate with guests via email or text messages to confirm reservations, provide check-in instructions, and collect feedback. This communication helps streamline the check-in process and keeps guests informed.

## 3. Proposed solution to the problem

The proposed solution to the problem involves the utilization of machine learning algorithms, different data points and various parameters like the number of days the guest is about to stay, the deposit amount paid, the type of room they reserved etc to predict whether the person who made the booking checks in or cancels. This predictive model can be used by the hotel management to make data-driven decisions, leading to better allocation of their resources and enhancing customer service.

- **Deposit or Prepayment:**

• With the help of this the predicted guests who are more likely to cancel make a non-refundable deposit or prepayment when booking a room. This ensures that the hotel has some revenue secured, even if the guest cancels at the last minute.

- **Discounted Non-Refundable Rates:**

• Offer discounted rates for non-refundable bookings. Guests who are certain about their plans can take advantage of lower prices, but they must accept that the booking is non-refundable

- **Overbooking Strategy:**

• With the help of this model we can Implement a controlled overbooking strategy, but exercise caution. Overbooking involves accepting more reservations than the hotel can accommodate and anticipating a certain percentage of no-shows and cancellations with the help of this model.

## 4. Reference to the problem

In terms of reference to these problems, it would be insightful to explore similar studies, industry reports, and research related to the field of hotel management for building an effective solution.

- **How hotel cancellation policies affect the search and booking decisions of deal-seeking customers**

  https://www.sciencedirect.com/science/article/abs/pii/S0278431910000320

- **Mass Cancellation Of Hotel Bookings In Himachal Severely Affects Tourism Business**

https://travelbizmonitor.com/mass-cancellation-of-hotel-bookings-in-himachal-severely-affects-tourism-business/

- **The travel industry faces a cancellation conundrum**

https://www.travelweekly.com/Middle-East-Africa-Travel/Insights/The-industry-faces-a-cancellation-conundrum

- **Three Most Common Trends Impacting Cancellation Rates**

https://revenue-hub.com/three-most-common-trends-impacting-cancellation-rates/

greatlearning
Learning for Life

# 4. CRITICAL ASSESSMENT OF TOPIC SURVEY

1. **Find the key area, gaps identified in the topic survey where the project can add value to the customers and business**

In the survey analysis, several key areas of opportunity and identified gaps have emerged where the project can significantly benefit both customers and the business. These include improving prediction accuracy, enabling real-time updates, enhancing interpretability, and ensuring robust data privacy measures to build trust.

- **Resource Management:**

Machine learning models can help hotels allocate resources such as housekeeping, front desk staff, and maintenance more effectively, reducing operational costs.

- **Data-Driven Decision-Making:**

Data-driven decisions help hotels adapt to market conditions, improve pricing strategies, and align services with customer preferences

- **Sustainability Efforts:**

Reducing resource waste associated with last-minute cancellations aligns with sustainability goals and reduces operational costs.

- **Marketing and Personalization:**

Machine learning models can analyze customer booking patterns to provide tailored promotions, increasing the likelihood of bookings and reducing cancellations.

2. **What key gaps are you trying to solve?**

The primary gaps the project aims to address involve enhancing the model's performance across various booking scenarios, enabling dynamic scalability to adapt to fluctuating booking volumes, providing explainable predictions for better decision-making, and implementing comprehensive data privacy measures to safeguard customer information.

- **Revenue Optimization:**

Predicting cancellations enables hotels to optimize their room inventory, reduce overbooking, and maximize revenue. It can also help identify opportunities to upsell or rebook canceled rooms.

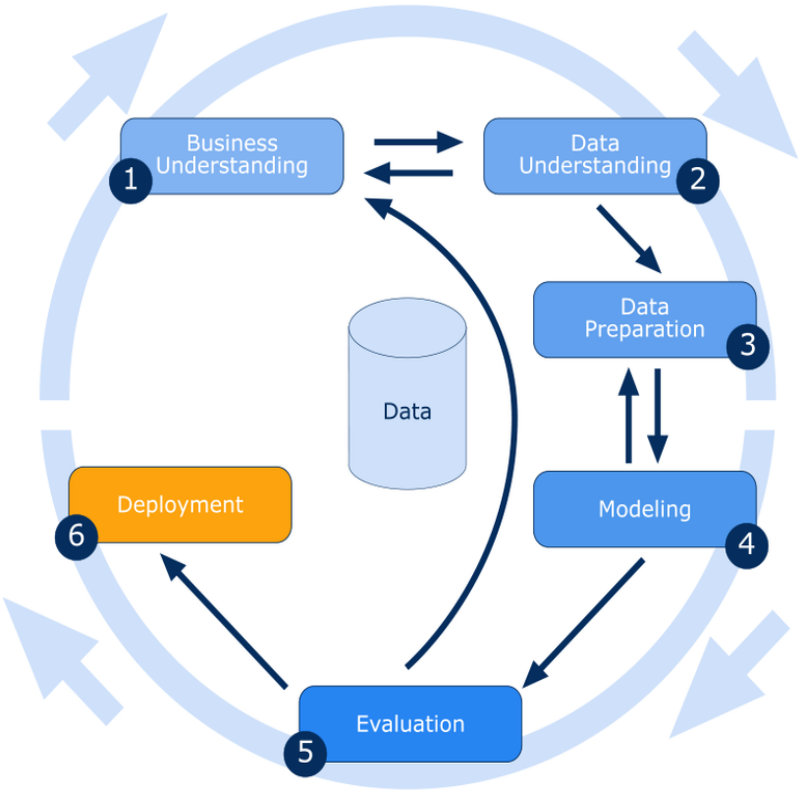
- **Risk Mitigation:**

Value for Business: Reducing last-minute cancellations minimizes the financial risk associated with empty rooms and the costs of rebooking guests.

- **Market Competitiveness:**

Hotels can gain a competitive edge by offering a more predictable and customer-centric booking experience.

# 5. METHODOLOGY TO BE FOLLOWED




1. **Business Understanding:**

The primary business objective of this project is to develop a predictive model that can accurately classify hotel booking reservations into two categories: those that are likely to be canceled and those that are not. This classification is essential for optimizing resource allocation and revenue management in the hotel industry.

2. **Data Understanding:**

1. Dataset: We utilized a hotel booking dataset containing various features such as booking details, customer information, and historical booking outcomes.

2. Data Exploration: We conducted an exploratory data analysis to understand the dataset's characteristics, distributions, and potential data quality issues.

3. **Data Preparation:**

1. Handling Missing Values: Imputing missing values for features with less than 40% missing data and dropped features with more than 40% missing data.

2. Feature Engineering: This step involves creating new features or transforming existing ones to enhance the predictive power of the dataset.

3. Data Splitting: The dataset will be divided into training, validation, and test sets to facilitate model development and evaluation.

**4. Modeling:**

1. Classification Algorithms: Model will be built with various binary classification algorithms, including logistic regression, decision trees, random forests, and support vector machines.

2. Hyperparameter Tuning: Model will be trained with different hyperparameters to improve model performance and generalization.

**5. Evaluation:**

The performance of each models evaluated using evaluation metrics such as accuracy, precision, recall, F1-score, ROC AUC and confusion matrix. The best-performing model will be selected based on the chosen evaluation metric, which aligned with business objectives.

greatlearning
Learning for Life

## 6. REFERENCES

| | |
|---|---|
| **Original owner of data** | **JESSE MOSTIPAK** |
| **Data set information** | This data set contains booking information for a city hotel and a resort hotel, and includes information such as when the booking was made, length of stay, the number of adults, children, and/or babies, and the number of available parking spaces, among other things. |
| **Link to web page** | https://www.kaggle.com/datasets/jessemostipak/hotel-booking-demand |

**********