iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 10 **Issue:** V **Month of publication:** May 2022

DOI: https://doi.org/10.22214/ijraset.2022.43036

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 10 Issue V May 2022- Available at www.ijraset.com

Hotel Booking Prediction using Machine Learning

Pranav Kumar[1], Sarthak Sharma[2]
[1, 2]*Department of Computer Science and Engineering, Raj Kumar Goel Institute of Technology, Delhi NCR, Ghaziabad 201206, UP, India*

Abstract: ***In this Project 'Hotel Booking Prediction', an accurate booking cancellation forecast by which user know the things related to hotel bookings very earlier. Booking cancellation has a significant effect on revenue which essentially affects request the board choices in the inn business. To reduce the cancellation effect, the hotel applies the cancellation model as the key to addressing this problem with the machine learning-based system developed. By combining data science tools and capabilities with human judgement and interpretation, this project aims to demonstrate how the predictive analysis of the model can contribute to synthesizing and predict about booking cancellation forecasting. Furthermore, this project aims, by detailing the full prediction & analysis, to give relaxation to user who want to apply in particular hotel. By Implement Various Algorithms like Logistic, KNN, Random Forest, Decision Tree, etc. to classify the data and also use Evaluation Matrix to separate categorical data in particular type, user can know the prediction up to the desired level. It prevents the hotel as well as Tourists to poor dealing of room. User/Customer have to enter certain field by which this model detects his prediction about the cancellation.***

Keywords: Logistic Regression, KNN, Random forest, Decision Tree, Evaluation Matrix.

I. INTRODUCTION

Hotels plays an important role for any person or traveler who are travelling from one destination to another. Hotel play an important role for tourists whether the tourist is local or international. Hotel provides many best services to the customer such as parking area, food, room service and also it provides services that is offered by customer. By providing these services Hotels take the valuable feedback from the customers. By these feedbacks Hotels maintains their reputation in the city/area. If the services are poor, the bookings of that hotel are low and if the services are awesome then high bookings in that hotel takes place.

In this model, the prediction possibility of a booking for a hotel based on different factors and also try to predict if they need special requests based on different features. In this project the dataset which we are using contains both International and Local Hotel data. Here we use many popular Machine Learning Algorithms like Decision Tree, Random Forest, KNN, Logistic Regression, etc. to predict the cancellation chances. This Model gives the prediction of Hotel Booking Cancellation up to the certain level of accuracy i.e. 95% (Approx.).

II. PROBLEM STATEMENT

1) Leverage Guest Data and Booking Behavior Patterns to devise a strategy for Hotel Booking Management, using Machine Learning.
2) With the increasing day to day trend of hotel cancellation at border time, it affects the local as well as international tourists more and more. Customer/Tourists don't have any idea about the pre-cancellation scenario of particular hotel at the time of booking of that hotel.

Solution: By 'HOTEL BOOKING PREDICTION' user can know about the cancellation percentage of Hotels in very early time. With the help of these predictions of Hotel Cancellation user can take decision about the booking of hotel/resort in Lead time or not, also it will allow hotel managers / revenue manager to take actions on bookings that's identified as "potentially going to be canceled", furthermore the development of these model should contribute to hotel revenue management.

III. LITERATURE SURVEY

A. Predicting Hotel Bookings Cancellation with a Machine Learning Classification

Booking cancellation is a very common thing in today's world, which can cause severe losses to the business owners. This paper describes how AI is used to identify which booking can be cancelled and prevent some losses. The machine learning model should be evaluated in the real time environment for accuracy. Prediction model of hotel booking cancellation no doubt the issue that can be resolved in the context of Design Science Research (DSR), as it need to develop an artifact, here in this particular case, a form of Revenue Management System (RMS), fulfilling the two requirements of DSR.

©IJRASET: All Rights are Reserved

B. Rising rate of Cancellation in the Hotel Industry

The growing trend of Hotels Industry is beneficial for Hotels but there are some problems too such as Rising Rate of Cancellation. The user cancels the booking of the hotel after seeing the reviews given by the people who booked the hotel already.

In Some case hotel owners treat the customer in bad way, this also affect the reputation as well as cancellations. The growing trend of Hotels Industry is beneficial for Hotels but there are some problems too such as Rising Rate of Cancellation. The user cancels the booking of the hotel after seeing the reviews given by the people who booked the hotel already. Now a days, people expecting a better accommodation at the Hotel site, if people found any lag in accommodation then they give poor rating of that Hotel. So, if we looking at percentage of cancellation then we found that the percentage of cancellation is increasing day by day. From a survey Cancellation rate rose from under 33% in 2014 to 40% in 2018. Also, during the COVID-19 pandemic this percentage gets increased because peoples book their Hotels in very earlier time and after changing situation day by day during pandemic and State Government implementing Lockdown in particular State, people sudden changes their plans and cancel the Hotel Bookings.

C. Aspect based Sentiment Oriented Summarization of Hotel Reviews

The reviews and the feedbacks of the customer play an important role in the image as well as the revenue system of the hotel. But most of the travelers don't read all reviews. The system analyzes the reviews and feedback by the customers. The feedbacks of the customer are gathered from the hotel's website and the stored as classes.

As per the study, the model analyzes the overlooked information by the customers and takes some essential steps. Finally, after processing all the data collected an emotional analysis is done. The hotels thereby can take the required steps to improve their service.

D. Application Of machine Learning in Hotel Industry A Critical Review

The growth in IT industry also affects the Hotel Industry. However, this change is quite slow. Many researchers are focused on testing and applying new artificial intelligence technology and learning equipment in the hotel industry. The study offers a brief knowledge about the use of Machine Learning and its combined technology in the hotel and tourism industry. Machine learning is quite trending these days.

IV. THE DATASET

The Dataset which is used in this project is taken from Kaggle.com. The Data is originally from the article Hotel Booking Demand Datasets, written by Nuno Antonio, Ana Almeida, and Luis Nunes for Data in Brief, Volume 22, February 2019.

This Dataset contains the details of Hotel situated in Local and International places. Most of the Hotel Data is from Portugal. Also, the data contains only two Hotels Name i.e. Resort and City Hotel assuming that the City Hotel means the Hotel Data for the City and Resort Hotel means the Hotels which are Hotel cum Resort. This Dataset contains four types of Data points, these are mentioned below.

1) Market Segments: Offline or Online sources from where the booking is generated and related variables.
2) Hotel & Revenue: City or Resort Hotel, ADR (Average Daily Rate), etc.
3) Customer Related: Variables describing the type of customer based on stay.
4) Cancellation History: Has there been cancellation been earlier? etc.

V. MACHINE LEARNING MODEL

The Machine Learning Model which is used in this project is built by applying various ML Algorithms like Decision Tree, Logistic Regression, KNN, Random Forest, XGBClassifier. The Algorithm which provides the highest accuracy is used in the final prediction of booking. For getting high accuracy in ML model, first the model gets trained using the basic training in which accuracy of different algorithms are listed below.

Evaluation Matrix					
	Logistic Regression	KNN	Decision Tree	Random Forest	XGB
Accuracy	0.804135	0.834391	0.831365	0.840316	0.832962
Recall	0.615210	0.744645	0.745665	0.763799	0.716083
Precision	0.810875	0.795592	0.788093	0.797043	0.811248
F1 Score	0.699620	0.769276	0.766292	0.780067	0.760701

Table – 1
Evaluation Matrix Of Differnet Algorithms

Here in this model "GridSearchCV" is used for Hyperparameter Tuning. In GridSearchCV approach, machine learning model is evaluated for a range of hyperparameter values. This approach is called GridSearchCV, because it searches for best set of hyperparameters from a grid of hyperparameters values.

Evaluation Matrix After Hyper Parameter Tuning					
	Logistic Regression	KNN	Decision Tree	Random Forest	XGB
Accuracy	0.803967	0.844182	0.837879	0.853679	0.850485
Recall	0.613510	0.745438	0.741244	0.759039	0.752465
Precision	0.811788	0.818137	0.805915	0.831616	0.828529
F1 Score	0.698857	0.780097	0.772228	0.793671	0.788667

Table – 2
Evaluation Matrix Of Differnet Algorithms After Hyper Parameter Tuning

VI- PROPOSED SYSTEM

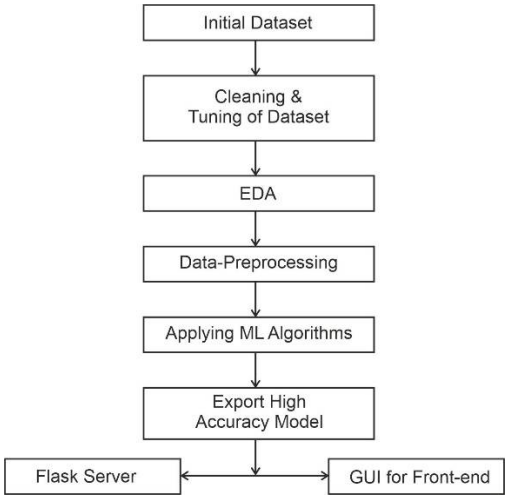


Fig. 1.: Proposed System of Hotel Booking Prediction

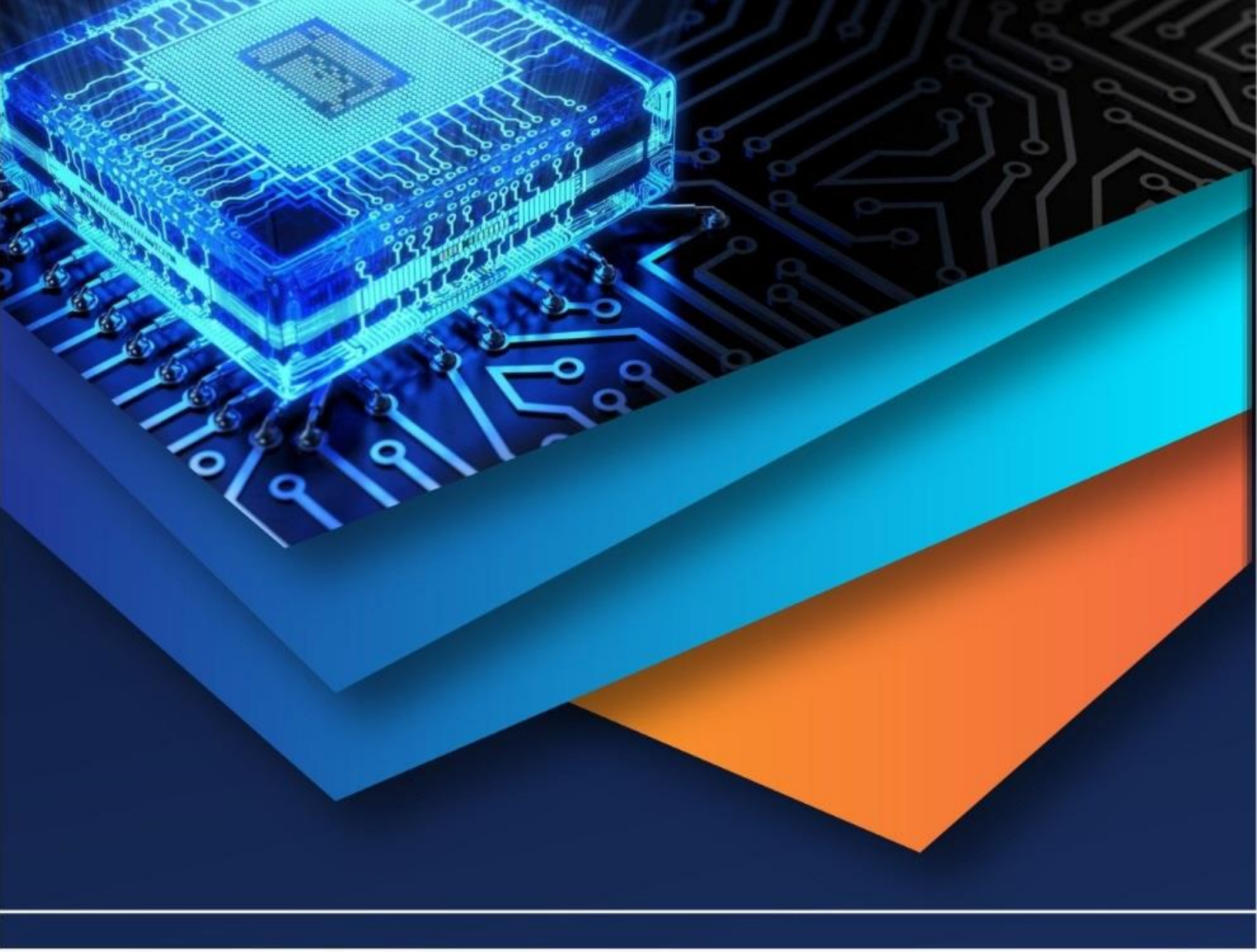
VII- CONCLUSIONS

Tuned Random Forest Has the Best Accuracy Among All Algorithm That We Tried from all the evaluation matrix to predict hotel cancellation classification case, we see that Tuned Random Forest has the best accuracy when it comes to predicting hotel cancellation based on certain features (85.2 %).

This model enables hotel managers to mitigate revenue loss derived from booking cancellations and to mitigate the risks associated with overbooking (reallocation costs, cash or service compensations, and, particularly important today, social reputation costs). This project also allows hotel managers to implement less rigid cancellation policies, without increasing uncertainty. This has the potential to translate into more sales, since less rigid cancellation policies generate more bookings.

REFERENCES

[1] Predicting Hotel Bookings Cancellation with machine learning classification model, Nuno Antonio, Ann de Almeida, Luis Nunes, IEEE International Conference on Machine Learning and Application, 2017.
[2] Application of Machine Learning in the Hotel Industry: A Critical Review, Dr. Eid Tourism, Archaeology Department, College of Arts, University of Hail, P.O.Box 2440Hail, Saudi Arabia, 2020.
[3] A Grouping Hotel Recommender System Based on Deep Learning and Sentiment Analysis, Fatemeh Abbasi, Ameneh Khadivar, Mohsen Yazdinejad, 2019.
[4] Machine Learning for Web Page Adaptation, Neetu Narwal, Dr. Sanjay Kumar Sharma, 2016.
[5] Studying the cancellation behaviour of the guests, Markku Vieru , 2016
[6] https://medium.com/analytics-steps/defining-predictive-modeling-in-machine-learning-887c23b7a278



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)