HOTEL RESERVATION SYSTEM USING MACHINE LEARNING

K Hannah Jessica - 20BCE7468

ABSTRACT

Our aim for the project is to develop a Hotel Reservation System which utilizes machine learning techniques to record, also manage detailed data of unlimited guests. The system will provide various functionalities to both users and hotel receptionists, such as making reservations, room cancellation, guest registration, total billing and room types. The project purpose is to show how the preliminary analysis can be used to synthesize and predict the reservation, and provide accurate predictions and analysis to give users confidence in their hotel selection. To achieve this, various algorithms such as Logistic Regression, KNN, Decision Tree, XG Boost, Candidate Elimination, Random Forest will be implemented to analyze data, also Evaluation Matrices will be used to separate categorical data into specific types. Users can obtain predictions up to the desired level, preventing poor management of rooms by hotels and providing a better experience for tourists.

KEYWORDS

Decision Tree, Random Forest, Logistic Regression, XG Boost, Candidate Elimination, Cancellation, forecasting, feedback, recall, F1-score, precision.

I. INTRODUCTION

Hotels have been essential part of all people's journey, whether they are traveling for pleasure or business. They will offer a range of services, such as food, parking and room service. Hotel management actively seek for a good feedback from customers to improve their services and maintain their reputation. If the quality of services offered by a hotel is poor, it can lead to a cancellation of reservations, while satisfactory service can result in high number of reservations.

In this project, our main motive is to predict hotel reservation cancellations based on various factors. They also identify if the customers have made any special requests. The dataset used in this project consists of both local and international hotel data. We have used possible machine learning algorithm such as Decision Tree, Candidate Elimination, Random Forest, Logistic Regression, XG Boost, KNN for predicting the accuracy for cancellation of reservations.

The goal of this project is to provide an accurate prediction of hotel reservation cancellations approximately 90% accuracy. Hotels can take protective measures to ensure that they have enough rooms available and avoid overbooking, which can lead to customer dissatisfaction. It will also assist in forecasting revenue and assist in better resource allocation. With machine learning techniques, hotels can make informed decisions and improve the overall guest experience. This model can be extended to other industries as well, supplying profitable insights to the customer and helping companies to adapt their services to encounter the customer' s requirements.

II. LITERATURE REVIEW

A. Machine Learning Classification used for Predicting Hotel Reservation

The primary objective for this paper is that to give a solution to all the difficulties loaded by the hotels during managing their reservation system by utilizing machine learning techniques. The traditional manual system of recording reservation details and guest information in books and files is time-consuming, inefficient and prone to errors. In contrast, the proposed hotel reservation system employs advanced machine learning algorithms to train and test the dataset, ensuring real-time evaluation of the model for accuracy.

The system's primary goal is to deliver an easy reservation process for hotel guests with features such as reservation or booking, cancellation of rooms, registration of guests, sign in and room management. With this system, hotels can secure their data better and record more detailed information about an unlimited number of guests. By doing so, it eliminates the inefficiencies of the manual system, and the booking process can be done faster and more efficiently, saving valuable time for both guests and receptionists.

The proposed system offers numerous benefits to hotels, including faster and more efficient booking processes, error-free data recording, and improved customer experience. Moreover, it can help hotels to manage their bookings better, increase their profitability and improve their reputation in the market. Overall, this system is a reliable and efficient solution for hotels to manage their reservation process and can be easily integrated into their existing infrastructure.

B. Increasing rate of Cancellation of the Hotel reservations

The hospitality industry is growing rapidly and while it offers many benefits to hotels, it also presents some challenges. One of the major issues faced by hotels is the rising rate of cancellations. Guests often cancel their bookings after reading reviews left by other guests or experiencing poor treatment from hotel staff. Nowadays, guests have high expectations for hotel accommodations, and if they encounter any issues, they tend to give poor ratings which affect the hotel's reputation and lead to more cancellations. According to article, the cancellation rate has increased in 2014 from 33% to 2018 of 40%. Moreover, during COVID-19 pandemic, the percentage of cancellations elevated as people made early reservations but had to cancel due to the changing situation and lockdowns implemented by state governments. The cancellation rate can significantly impact a hotel's revenue and reputation. Therefore, it is crucial for hotels to address this issue by using advanced technologies such as machine learning to predict the likelihood of cancellations and take proactive measures to mitigate them. By doing so, hotels can provide better services to guests, improve their reputation, and increase revenue.

C. Sentiment Oriented Summarization of Hotel Feedback

Reviews and feedback provided by customers have a significant impact on the image and revenue of a hotel. However, many travellers may not read all the reviews, making it challenging for hotels to gauge their performance accurately. To address this issue, a system can analyze customer feedback and reviews collected from the hotel's website and classify them according to various categories. By analyzing this data, hotels can identify areas for improvement and take necessary steps to enhance their services and customer satisfaction.

In recent years, the hotel industry has been slow in adopting technological advancements compared to other industries. However, with the growth of the IT industry, diverse investigators are focusing to test and to implement new AI techniques and learning equipments for the hotel industries. Machine learning is one such technology that has gained significant attention in the hotel and tourism industry. The study offers valuable insights is applied for machine learning, its benefits, and its potential to transform the industry.

III. PROPOSED METHOD

This Dataset consists of 36,275 rows and 19 columns which is taken from Kaggle. The research paper used is Hotel Booking System for Ridel Hotel Kota Bharu, written by Nurul Khairiyah Rahimi, Azizul Azhar Ramli, May 2022.

The Models of Machine Learning used in this project are developed by applying ML Algorithms like Decision Tree, KNN, Logistic Regression, Candidate Elimination, Random Forest, XG Boost, Gradient Boosting. The Algorithm with the highest accuracy is considered as the ultimate prediction of reservation. To get the highest accuracy for a given ML model, firstly the model will get trained using the algorithms for which the accuracy is predicted which are listed below.

Table – 1 Evaluation Matrix Before Prediction

Evaluation Matrix					
Algorithms	Logistic Regression	Random Forest	Decision Tree	KNN	XG BOOST
Accuracy	0.804135	0.881316	0.880365	0.7918	0.890839

Evaluation Matrix Before Prediction			
Algorithms	Logistic Regression	Random Forest	Decision Tree
Recall	0.615210	0. 743799	0. 795665
Precision	0.740875	0. 867043	0. 832093
F1 Score	0.675620	0.800067	0.806292

Table – 2 Evaluation Matrix After Prediction

Evaluation Matrix					
Algorithms	Logistic Regression	Decision Tree	Random Forest	KNN	XG BOOST
Accuracy	0.803967	0.883879	0.887679	0.802195	0.880817

Evaluation Matrix After Tuning			
Algorithms	Logistic Regression	Decision Tree	Random Forest
Recall	0.613510	0.791244	0.749039
Precision	0.741788	0.835915	0.871616
F1 Score	0.678857	0.812228	0.803671

The main motive of this project is to detect the accuracy of the booking status of the reservation for a particular hotel. The dataset used is Hotel reservation consisting of 36,275 rows and 19 columns. Firstly, Data preparation is used for Exploratory data analysis (EDA) where we understand and visualize about the data we are working on, then data preprocessing technique is used to import all the libraries required for training and testing this dataset. After importing the dataset, we start data cleaning where we remove all the incorrect or corrupted data and fix the data. This process will reduce errors and also improves data quality.

After Data preprocessing, label encoding and standard scaler are used for the algorithms. Split the data and train them as the dataset we have taken is a supervised learning of classification we have to validate and test it. The algorithms we have used for this project is Decision Tree, Random Forest, Logistic Regression, Candidate Elimination, KNN and XG Boost. Here, we get the most accuracy for Random Forest of nearly 89% after tuning.

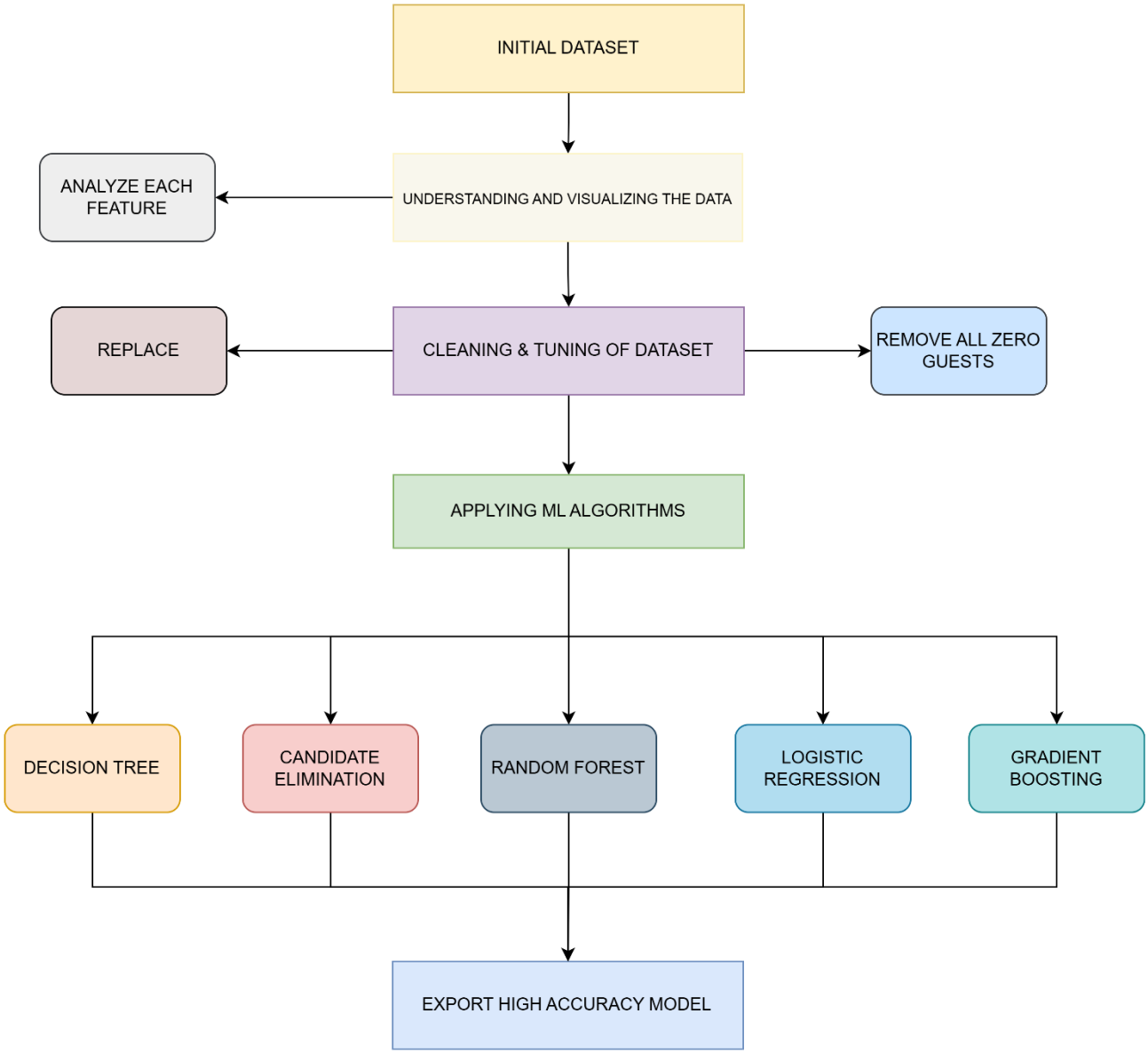

Fig. 1: PROPOSED SYSTEM FOR HOTEL RESERVATION SYSTEM

IV. RESULTS & DISCUSSIONS

The main motive for this project is to predict accuracy of the booking status of the hotel. The Random Forest modelling present best classification accuracy of 89%. It means that most of the people are satisfied with room booking and their services. Comparing with the existing project, I have got the better result as the existing model accuracy was 84% for Random Forest Model. From the preliminary analysis for the model, precision success rate is of 87.16%, recall

success rate is 74.39% and F1 success score is 80%. The forecasting reveals that the result values obtained are best as it is performing better than the existing project. Although there are few limitations but those can be improved in the future scopes.

V. CONCLUSION & FUTURE SCOPES

CONCLUSION

After experimenting with various machine learning algorithms to anticipate the classification of hotel reservations, and we found that the Random Forest algorithm has the best and highest accuracy among all the evaluation metrics, with an accuracy of 88.7%. This is a significant improvement for hotel managers as it allows them to minimize the loss of cancellations of booking and reduce the risks related to overbooking. Overbooking can result in reallocation costs, cash or service compensations, and damage to the hotel's social reputation, which is particularly crucial in today's digital age. By implementing the Random Forest model, hotel managers can also offer less rigid cancellation policies without increasing uncertainty, which may lead to increased sales as guests are more likely to book rooms. This project provides a valuable tool for hotel managers to optimize their reservation system and maximize their profits while providing a better guest experience. We conclude that this project has got better accuracy than the existing ones so it can be used by all.

FUTURE SCOPES

- We can improve the accuracy of the model by testing them using the other techniques.
- We can create an app for the people to be more convenient for them to use it.
- In this project we can try using the other algorithms as it is considered under classification.
- We can make the payment section where they can pay it easily and conveniently so that it may not get failed.

VI. REFERENCES

[1] A. J. Sánchez-Medina and C. Eleazar, "Using machine learning and big data for efficient forecasting of hotel booking cancellations", *International Journal of Hospitality Management*, vol. 89, pp. 102546, 2020.

[2] D.R. Morales, J. Wang, Forecasting cancellation rates for services booking revenue management using data mining, Eur. J. Oper. Res., 202 (2010), pp. 554-562

[3] Q. Ye *et al.*, The impact of online user reviews on hotel room sales
International Journal of Hospitality Management (2009)

[4] M. Falk, M. Vieru, Modelling the cancellation behaviour of hotel guests

Int. J. Contemp. Hosp. M, 30 (2018), pp. 3100-3116

[5] Ladhari R, Michaud M (2015) eWOM effects on hotel booking intentions, attitudes, trust, and website perceptions. Int J Hosp Manag 46:36–45

[6] N. Antonio, A. De Almeida, L. Nunes, Hotel booking demand datasets, Data Brief, 22 (2019), pp. 41-49

[7] Chang Y-C, Ku C-H, Chen C-H (2019) Social media analytics: Extracting and visualizing Hilton hotel ratings and reviews from TripAdvisor. Int J Inf Manage 48:263–279

[8] G.X. Gao, J.W. Bi, Hotel booking through online travel agency: optimal Stackelberg strategies under customer-centric payment service, Ann. Tour. Res., 86 (2021), Article 103074

[9] Zhang Z, Ye Q, Law R (2011) Determinants of hotel room price: an exploration of travelers' hierarchy of accommodation needs. Int J Contemporary Hospitality Manag

[10] Hu N, Zhang T, Gao B, Bose I. What do hotel customers complain about? Text analysis using structural topic model. Tourism Management. 2019;72:417-426.

[11] Bitran, G.R., Gilbert, S.M., 1996. Managing hotel reservations with uncertain arrivals. Oper. Res. 44, 35–49.

[12] Barsky JD (1992) Customer satisfaction in the hotel industry: Meaning and measurement. Hospitality Res J 16(1):51–73

[13] Ahani A et al (2019) Revealing customers' satisfaction and preferences through online review analysis: the case of Canary Islands hotels. J Retail Consum Serv 51:331–343

[14] A.J. Sánchez-Medina, E.C. Sánchez, Using machine learning and big data for efficient forecasting of hotel booking cancellations

Int. J. Hosp. Manag., 89 (2020), Article 102546

[15] J. Novakovic and S. Turina, "Hotel reservation cancellations: analysis and prediction using machine learning algorithms", *ACADEMIC JOURNAL*, pp. 4.