

HOTEL BOOKING MANAGEMENT

¹PALLIKONDA SAROJA, ²B. MOUNICA

¹Student, ²Associate Professor

¹²DEPARTMENT OF COMPUTER APPLICATIONS

¹²RAJAHMAHENDRI INSTITUTE OF ENGINEERING AND TECHNOLOGY, Rajahmahendry

Abstract—Last-minute hotel reservation cancellations, often due to schedule changes, significantly impact hotel revenue. We can predict these cancellations by analyzing booking parameters such as the number of guests, rooms booked, and booking duration. Our study employed four models: Logistic Regression, SVM (Linear and RBF kernels), Decision Tree, and Random Forest. We evaluated these models using training and test datasets, with the Random Forest model achieving the highest accuracy at 99%. Additionally, we introduced a weighted average technique to help users visualize model performance and select the most suitable model for their needs.

Index Terms—Hotel Booking, Management

I. Introduction

Tourism is a crucial global industry, with foreign visit or growing from 500 million in 1995 to 1400 million in 2018. In 2018 alone, there was a 5% increase in tourists, with export earnings reaching USD 1.7 trillion [1, 2, 3]. This industry drives economic growth and creates jobs. New technologies in virtual reservations have transformed consumer behaviour. Hotel cancellations impact revenue through lost income, higher advertising costs, and reduced profit margins. This study examines hotel cancellation policies, noting that hotels suffer when they cannot resell rooms, pay more for advertising, lower rates to resell at the last minute, and must reevaluate guest arrangements.

The rise in cancellations necessitates a machine-learning system to predict which reservations will most likely be canceled. INN Hotels Group in Portugal requires a data-driven solution to address this issue. Analyzing data to identify critical variables can help develop predictive models, allowing hotels to anticipate cancellations and create effective cancellation and refund policies. Accurate cancellation prediction is crucial in modern booking systems to avoid significant disruptions and inefficiencies caused by inaccurate forecasts. The complexity of the current booking landscape and increased demand for lodging make manual identification of true cancellations challenging. The contributions of the paper are as follows:

- i) The dataset has been analyzed using the Exploratory Data Analysis (EDA) technique to determine the factor (s) for cancellation.
- ii) We have created a prediction model for hotel booking cancellation using four machine learning algorithms.
- iii) Recall, accuracy, precision, and f1 score are used to asses show well the machine learning models perform. A comparative performance of the suggested models is also provided.
- iv) Lastly, the trained ensemble model has been integrated into a web application with an interface that allows users to check a new hotel booking request and analyze it with different visual analytics for possible cancellations.

II. SYSTEM ANALYSIS

EXISTING SYSTEM

The existing hotel reservation systems are often reactive, lacking predictive insights. They rely heavily on traditional rules or manual observations to manage booking cancellations, with no robust machine learning-based approach to predict cancellations. Some hotels use simple data from past cancellations, but this data is generally not analyzed effectively to predict future cancellation trends.

DISADVANTAGES OF EXISTING SYSTEM

- **Limited Predictive Capabilities:** The absence of machine learning models means cancellations are managed based on historical trends rather than predictions, leading to reactive rather than proactive management.
- **Revenue Loss:** Last-minute cancellations impact revenue as they reduce occupancy rates, and rebooking on short notice often requires discounts.
- **Higher Costs:** Hotels incur additional costs for advertising and adjusting prices to fill last-minute cancellations.
- **Customer Service Challenges:** Without predictive insights, managing room availability for loyal or high-value customers becomes more complex.

PROPOSED SYSTEM

The proposed system uses machine learning algorithms to predict the likelihood of booking cancellations. It employs multiple models, including Logistic Regression, SVM (Support Vector Machine), Decision Tree, and Random Forest, with a focus on achieving high accuracy in predicting cancellations. Random Forest was identified as the most effective model in the study, achieving a prediction accuracy of 99%.

The proposed system also includes a web-based visualization tool for hotel managers. This tool enables users to analyze booking patterns and predictions, aiding in informed decision-making and proactive management of potential cancellations.

ADVANTAGES OF PROPOSED SYSTEM

- **Improved Prediction Accuracy:** The Random Forest model achieved 99% accuracy, enabling more reliable predictions.
- **Revenue Optimization:** By anticipating cancellations, hotels can proactively manage bookings, minimize losses, and improve occupancy rates.
- **Cost Savings:** Reduces last-minute advertising and discounting costs by accurately identifying high-risk cancellations in advance.
- **Enhanced Customer Service:** Allows hotels to maintain availability for loyal customers and manage cancellations more effectively.
- **User-Friendly Visualization:** The integrated web-based tool provides accessible insights and customizable analytics for hotel managers.

Existing System Algorithm

The existing system primarily relies on heuristic or rule-based approaches without sophisticated algorithms. Rules are often derived from historical patterns, such as cancellation frequency or booking lead times, but these are not backed by robust predictive models.

Proposed System Algorithm

The proposed system uses several machine learning algorithms, with the following details:

- **Logistic Regression:** Used as a baseline model to classify cancellations but suffers from lower accuracy in complex cases.
- **Support Vector Machine (SVM):** Utilizes both Linear and RBF kernels for improved classification but still has limitations in recall for cancellations.
- **Decision Tree:** Provides interpretable predictions but is prone to overfitting.
- **Random Forest:** The most effective model in the study, combining multiple decision trees to enhance prediction accuracy and generalization, achieving a high F1 score and reduced false positives/negatives.

FEASIBILITY STUDY

The feasibility of the project is analyzed in this phase and business proposal is put forth with a very general plan for the project and some cost estimates. During system analysis the feasibility study of the proposed system is to be carried out. This is to ensure that the proposed system is not a burden to the company. For feasibility analysis, some understanding of the major requirements for the system is essential.

Three key considerations involved in the feasibility analysis are,

- ◆ ECONOMICAL FEASIBILITY
- ◆ TECHNICAL FEASIBILITY
- ◆ SOCIAL FEASIBILITY

ECONOMICAL FEASIBILITY

This study is carried out to check the economic impact that the system will have on the organization. The amount of fund that the company can pour into the research and development of the system is limited. The expenditures must be justified. Thus the developed system as well within the budget and this was achieved because most of the technologies used are freely available. Only the customized products had to be purchased.

TECHNICAL FEASIBILITY

This study is carried out to check the technical feasibility, that is, the technical requirements of the system. Any system developed must not have a high demand on the available technical resources. This will lead to high demands on the available technical resources. This will lead to high demands being placed on the client. The developed system must have a modest requirement, as only minimal or null changes are required for implementing this system.

SOCIAL FEASIBILITY

The aspect of study is to check the level of acceptance of the system by the user. This includes the process of training the user to use the system efficiently. The user must not feel threatened by the system, instead must accept it as a necessity. The level of acceptance by the users solely depends on the methods that are employed to educate the user about the system and to make him familiar with it. His level of confidence must be raised so that he is also able to make some constructive criticism, which is welcomed, as he is the final user of the system.

SOFTWARE REQUIREMENTS & SPECIFICATIONS

REQUIREMENT ANALYSIS

The project involved analyzing the design of few applications so as to make the application more users friendly. To do so, it was really important to keep the navigations from one screen to the other well ordered and at the same time reducing the amount of typing the user needs to do. In order to make the application more accessible, the browser version had to be chosen so that it is compatible with most of the Browsers.

III. SYSTEM DESIGN

SYSTEM ARCHITECTURE:

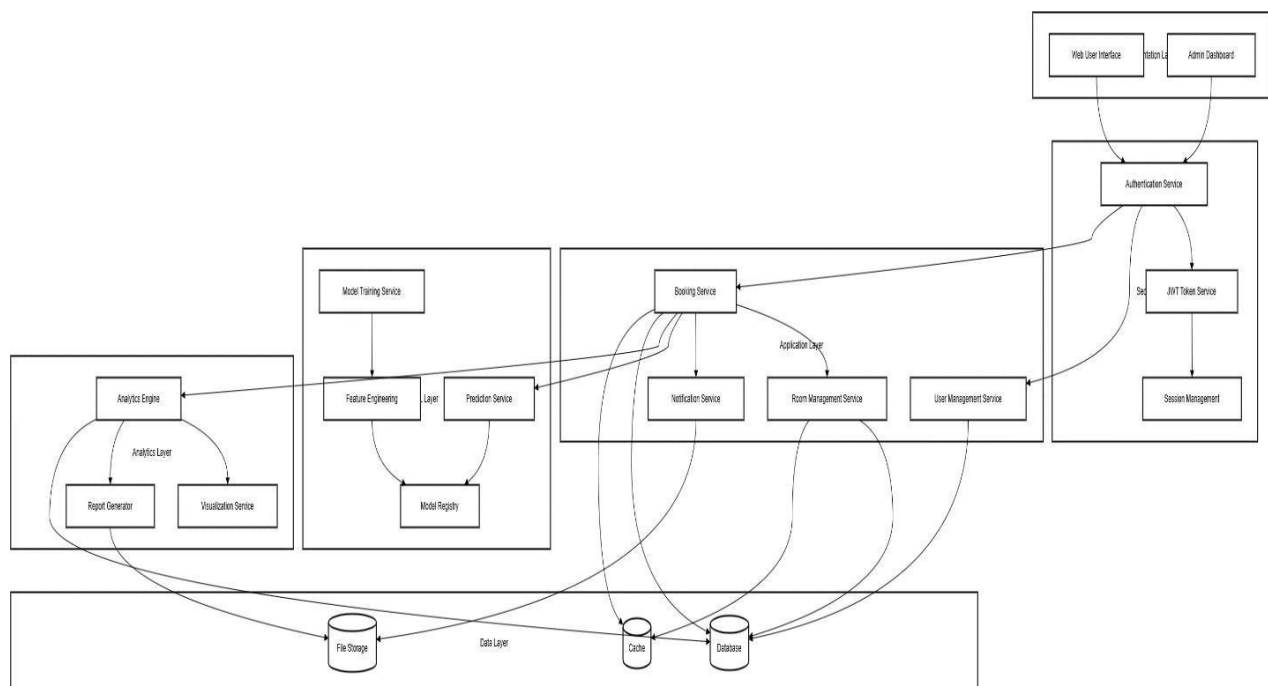


Fig-3.1: System Architecture

INPUT DESIGN

The input design is the link between the information system and the user. It comprises the developing specification and procedures for data preparation and those steps are necessary to put transaction data in to a usable form for processing can be achieved by inspecting the computer to read data from a written or printed document or it can occur by having people keying the data directly into the system. The design of input focuses on controlling the amount of input required, controlling the errors, avoiding delay, avoiding extra steps and keeping the process simple. The input is designed in such a way so that it provides security and ease of use with retaining the privacy. Input Design considered the following things:

- What data should be given as input?
- How the data should be arranged or coded?
- The dialog to guide the operating personnel in providing input.

- Methods for preparing input validations and steps to follow when error occur.

OBJECTIVES

1. Input Design is the process of converting a user-oriented description of the input into a computer-based system. This design is important to avoid errors in the data input process and show the correct direction to the management for getting correct information from the computerized system.
2. It is achieved by creating user-friendly screens for the data entry to handle large volume of data. The goal of designing input is to make data entry easier and to be free from errors. The data entry screen is designed in such a way that all the data manipulates can be performed. It also provides record viewing facilities.
3. When the data is entered it will check for its validity. Data can be entered with the help of screens. Appropriate messages are provided as when needed so that the user will not be in maize of instant. Thus the objective of input design is to create an input layout that is easy to follow

OUTPUT DESIGN

A quality output is one, which meets the requirements of the end user and presents the information clearly. In any system results of processing are communicated to the users and to other system through outputs. In output design it is determined how the information is to be displaced for immediate need and also the hard copy output. It is the most important and direct source information to the user. Efficient and intelligent output design improves the system's relationship to help user decision-making.

1. Designing computer output should proceed in an organized, well thought out manner; the right output must be developed while ensuring that each output element is designed so that people will find the system can use easily and effectively. When analysis design computer output, they should Identify the specific output that is needed to meet the requirements.
2. Select methods for presenting information.
3. Create document, report, or other formats that contain information produced by the system.

The output form of an information system should accomplish one or more of the following objectives.

- Convey information about past activities, current status or projections of the
- Future.
- Signal important events, opportunities, problems, or warnings.
- Trigger an action.
- Confirm an action.

DATA FLOW DIAGRAM:

1. The DFD is also called as bubble chart. It is a simple graphical formalism that can be used to represent a system in terms of input data to the system, various processing carried out on this data, and the output data is generated by this system.
2. The data flow diagram (DFD) is one of the most important modeling tools. It is used to model the system components. These components are the system process, the data used by the process, an external entity that interacts with the system and the information flows in the system.
3. DFD shows how the information moves through the system and how it is modified by a series of transformations. It is a graphical technique that depicts information flow and the transformations that are applied as data moves from input to output.

4. DFD is also known as bubble chart. A DFD may be used to represent a system at any level of abstraction. DFD may be partitioned into levels that represent increasing information flow and functional detail.

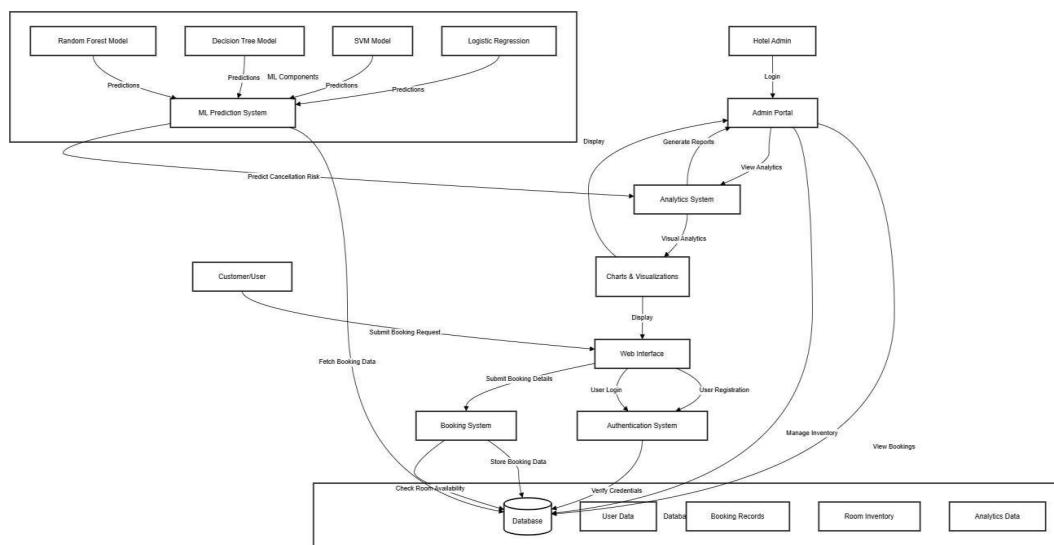


Fig-3.2: Data Flow Diagram

UML DIAGRAMS

UML stands for Unified Modeling Language. UML is a standardized general-purpose modeling language in the field of object-oriented software engineering. The standard is managed, and was created by, the Object Management Group.

The goal is for UML to become a common language for creating models of object oriented computer software. In its current form UML is comprised of two major components: a Meta- model and a notation. In the future, some form of method or process may also be added to; or associated with, UML.

The Unified Modeling Language is a standard language for specifying, Visualization, Constructing and documenting the artifacts of software system, as well as for business modeling and other non-software systems.

The UML represents a collection of best engineering practices that have proven successful in the modeling of large and complex systems.

The UML is a very important part of developing objects oriented software and the software development process. The UML uses mostly graphical notations to express the design of software projects.

GOALS:

The Primary goals in the design of the UML are as follows:

1. Provide users a ready-to-use, expressive visual modeling Language so that they can develop and exchange meaningful models.
2. Provide extendibility and specialization mechanisms to extend the core concepts.
3. Be independent of particular programming languages and development process.
4. Provide a formal basis for understanding the modeling language.

5. Encourage the growth of OO tools market.
6. Support higher level development concepts such as collaborations, frameworks, patterns and components.
7. Integrate best practices.

IV. CONCLUSIONS

By reviewing reservation and cancellation history, businesses can reduce cancellations. Implementing a user friendly online platform enhances customer satisfaction. Analysis shows distinct customer segments, allowing targeted marketing. Customers with longer lead times have higher cancellation rates, highlighting the importance of competitive pricing. Optimizing pricing by analyzing market trends and competitors' strategies can help. Implementing dynamic pricing based on lead time and seasonal demand, offering non-refundable rates with discounts, and personalized email campaigns can reduce cancellations, increase bookings, and boost customer satisfaction and loyalty.

References

- [1] Eleazar C-Sánchez Agustín, J. Sánchez-Medina, Laura Romero- Domínguez: Forecasting Hotel-booking Cancellations Using Personal Name Records: An Artificial Intelligence Approach (2022),doi:10.1007/978-981-16-9268-0_1
- [2] Data Statistics of World Tourism Organization, https://view.officeapps.live.com/op/view.aspx?src=https%3A%2F%2Fwebunwto.s3.eu-west-1.amazonaws.com%2Fs3fs-public%2F202303%2Funwto-all-datadownload_032023.xlsx&wdOrigin=BROWSELINK
- [3] Tourism Dashboard of World Tourism Organization, <https://www.unwto.org/tourism-data/unwto-tourism-dashboard>
- [4] Nuno Antonio, Ana Maria De Almeida, Luís Nunes: Predicting hotel booking cancellations to decrease uncertainty and increase revenue (2017),doi: 10.18089/tms.2017.13203
- [5] Nuno Antonio, Ana de Almeida, Luis Nunes: Predicting Hotel Bookings Cancellation with a Machine Learning Classification Model (2017), doi:10.1109/ICMLA.2017.00-11