

# Enhancing Hospitality Management Through ML-Based Cancellation Prediction

Miriyala Akash, department of Computer Science and Engineering, GNITC, 22-5H9, [22wj1a05h9@gniindia.org](mailto:22wj1a05h9@gniindia.org)

Masa Yashwanth, department of Computer Science and Engineering, GNITC, 22-5H6, [22wj1a05h6@gniindia.org](mailto:22wj1a05h6@gniindia.org)

Kavali Bhanusree, department of Computer Science and Engineering, GNITC, 23-515, [23wj5a0515@gniindia.org](mailto:23wj5a0515@gniindia.org)

Dr. MVA Naidu, Assistant Professor, department of Computer Science and Engineering, GNITC,

\*\*\*

**Abstract** - Hotel booking cancellations create significant challenges for the hospitality industry by affecting revenue management, demand forecasting, and resource allocation. Traditional prediction methods often fail to capture the complex behavioral patterns associated with customer cancellations. This research proposes a machine learning-based approach using a Multi-Layer Perceptron (MLP) classifier to predict hotel booking cancellations accurately. The model analyzes various booking attributes such as lead time, room type, customer segment, location, and booking history to detect patterns associated with cancellation behavior. Data preprocessing techniques including normalization, handling missing values, and encoding categorical variables are applied to improve model performance. Hyperparameter tuning is performed to optimize hidden layers, activation functions, and learning rates. Experimental results demonstrate that the proposed model provides high prediction accuracy and supports hotel managers in making informed decisions related to demand forecasting, overbooking strategies, and revenue optimization. The system ultimately contributes to improved operational efficiency and customer satisfaction in the hospitality sector. Hotel booking cancellations create major challenges for the hospitality industry by affecting revenue management and demand forecasting. This research proposes a machine learning approach using a **Multi-Layer Perceptron (MLP)** classifier to predict hotel booking cancellations based on historical booking data. The model helps hotel managers improve decision-making, optimize resources, and reduce financial losses caused by cancellations.

**Key Words:** Machine Learning, Hotel Booking Cancellation, Multi-Layer Perceptron, Hospitality Management, Predictive Analytics.

## 1. INTRODUCTION

The hospitality industry relies heavily on accurate demand forecasting to manage room occupancy, staffing, and revenue optimization. However, hotel booking cancellations have become a significant issue that disrupts operational planning and leads to financial losses. Unexpected cancellations may cause unoccupied rooms, inefficient resource allocation, and inaccurate demand predictions.

Traditional approaches for managing cancellations rely mainly on historical averages and statistical methods. These approaches are often limited in their ability to capture complex behavioral patterns associated with customer booking behavior. With the growth of data-driven technologies, machine learning and artificial intelligence techniques have gained attention for solving such predictive problems.

Machine learning models can analyze large volumes of historical booking data to identify patterns and relationships that influence cancellation behavior. Among these models, the

Multi-Layer Perceptron (MLP) classifier is particularly effective in capturing non-linear relationships between variables.

In this research, an MLP-based predictive model is developed to forecast hotel booking cancellations. By analyzing customer and booking-related attributes such as lead time, booking channel, room type, and customer segment, the proposed system provides accurate predictions and valuable insights for hotel managers. This enables improved decision-making in areas such as overbooking policies, pricing strategies, and demand forecasting.

## 2. LITERATURE REVIEW

In recent years, researchers have increasingly explored the use of **machine learning, deep learning, and big data techniques** to improve forecasting and decision-making in the hospitality industry. Several studies have focused on predicting hotel demand, customer behavior, booking cancellations, and sales performance using advanced analytical methods.

Dong Zhang and Baozhuang Niu (2024) proposed a deep learning approach for hotel demand forecasting by leveraging signals extracted from online reviews. Their research highlights that traditional forecasting methods mainly rely on sentiment analysis, which may not capture all useful signals from customer reviews. To address this limitation, they introduced a framework based on **Systematic Functional Linguistics (SFL)** theory that extracts ideational, textual, and interpersonal signals from review data. Their model demonstrates improved forecasting performance by analyzing complex relationships in multivariate time-series data.

Gang Chen, Lihua Huang, Shuaiyong Xiao, Chenghong Zhang, and Huimin Zhao (2024) developed a novel deep learning method to enhance sales prediction using **multimodal online reviews**. Their study introduces the concept of **customer attention**, which reflects how much influence certain reviews have on customer purchasing behavior. The researchers proposed four indicators of customer attention: **timeliness, semantic diversity, voting awareness, and multimodal interaction**. Using neural network attention mechanisms, their model achieved higher prediction accuracy compared to traditional deep learning models when forecasting hotel sales and occupancy rates.

Ismael Gomez Talal, Lydia González Serrano, José Luis Rojo-Álvarez, and Pilar Talón Ballesteros (2024) investigated the use of **big data analytics and unsupervised machine learning** to reduce food waste in restaurants. Their system analyzes restaurant ticket data to identify relationships between customer purchases and product demand across different time periods. The model applies multicomponent analysis and bootstrap resampling to discover patterns such as product

combinations, peak demand times, and cross-selling opportunities. Their research demonstrates how machine learning can improve operational efficiency and support better decision-making in hospitality management.

Ismael Gómez-Talal, Mana Azizoltani, Pilar Talón-Ballester, and Ashok Singh (2024) focused specifically on predicting hotel booking cancellations using machine learning techniques. They proposed a predictive model based on **stacked generalization combined with Explainable Artificial Intelligence (XAI)**. Their approach integrates multiple predictive models, including linear, tree-based, and deep learning models, into a single meta-model. The system achieved a prediction accuracy of approximately **96%** while also providing interpretability by identifying key factors influencing cancellations, such as room type, hotel location, and customer segment.

Zhenkun Liu, Ping Jiang, Jianzhou Wang, and Zhiyuan Du (2023) proposed a **profit-driven hospitality order cancellation prediction model** using the Extreme Gradient Boosting (XGBoost) algorithm. Unlike traditional prediction models that focus only on accuracy, their approach emphasizes maximizing profit for hotel management. The model analyzes customer booking data and uses grid search to determine optimal hyperparameters for XGBoost. Their results show that the proposed method not only improves prediction accuracy but also increases hotel profitability by enabling more effective customer retention strategies.

Overall, these studies demonstrate that **machine learning and deep learning techniques play an important role in hospitality analytics**, particularly in forecasting demand, predicting cancellations, and improving operational efficiency. However, many existing models focus either on demand forecasting or profitability analysis rather than specifically predicting booking cancellations using neural network-based approaches. Therefore, this research proposes a **Multi-Layer Perceptron (MLP)-based machine learning model** to accurately predict hotel booking cancellations and support better decision-making in hotel management.

### 3. RELATED WORK

Recent advancements in **machine learning, artificial intelligence, and data analytics** have significantly influenced research in the hospitality industry, particularly in areas such as demand forecasting, booking behavior analysis, and cancellation prediction. Researchers have explored various predictive models to improve operational efficiency and revenue management in hotels.

Liu et al. (2023) proposed a **profit-driven hospitality order cancellation prediction model** based on the Extreme Gradient Boosting (XGBoost) algorithm. Unlike traditional models that focus only on prediction accuracy, their approach emphasizes maximizing profit by identifying potential cancellations that may cause revenue loss. The model applies hyperparameter optimization techniques to determine the best model configuration. Experimental results showed that the proposed approach performs better than several baseline models in both profitability and prediction accuracy.

Zhang and Niu (2024) explored the use of **deep learning techniques combined with online review analysis** for hotel demand forecasting. Their research applies a linguistic-based framework to extract meaningful signals from customer

reviews. By analyzing textual information from reviews, the model captures customer perceptions and behavioral trends that influence hotel demand. The study demonstrates that integrating customer feedback into predictive models can significantly enhance forecasting accuracy.

Chen et al. (2024) introduced a **multimodal deep learning framework** for improving hotel sales prediction using customer-generated review data. Their research focuses on understanding customer attention and how it affects purchasing decisions. The proposed system extracts information from multiple modalities such as text, ratings, and review interactions. Using neural network attention mechanisms, the model captures complex relationships between these features and improves the accuracy of sales predictions in hospitality environments.

Gómez-Talal et al. (2024) presented an interpretable machine learning approach for predicting hotel booking cancellations. Their method combines multiple predictive models through **stacked generalization**, which integrates linear models, tree-based algorithms, and deep learning techniques. In addition to improving prediction performance, the researchers applied **Explainable Artificial Intelligence (XAI)** to identify key factors influencing cancellation behavior. Their study highlights that features such as customer segment, booking lead time, and room type play significant roles in predicting cancellations.

Another study by Gómez-Talal et al. (2024) investigated the application of **big data analytics and unsupervised machine learning** to analyze restaurant ticket data. The research focuses on identifying relationships between product demand and customer purchasing behavior. By applying clustering and statistical analysis techniques, the system helps restaurant managers make better decisions regarding product inventory, sales strategies, and resource management.

Despite these advancements, many existing approaches rely on complex ensemble models or focus primarily on demand forecasting and sales prediction rather than directly addressing cancellation prediction. Furthermore, some methods require large computational resources and complex architectures, making them difficult to implement in real-world hotel management systems.

To address these limitations, the proposed system introduces a **machine learning-based predictive model using a Multi-Layer Perceptron (MLP) neural network**. The model analyzes various booking attributes such as lead time, room type, customer segment, booking channel, and previous booking history to identify patterns associated with cancellations. Compared to traditional approaches, the proposed model aims to provide **high prediction accuracy, efficient training, and practical implementation for real-time hotel management systems**.

### 4. PROPOSED METHODOLOGY

The proposed methodology focuses on developing an intelligent system to predict hotel booking cancellations using a Multi-Layer Perceptron (MLP) Classifier. The system analyzes historical booking data and identifies patterns that influence cancellation behavior. The methodology consists of several stages including data acquisition, data preprocessing, model training, prediction, and performance evaluation.

#### 1. Data Acquisition

The first step in the methodology is collecting the dataset related to hotel bookings. The dataset contains historical reservation

records with multiple attributes that influence customer cancellation behavior.

The dataset includes features such as:

- Booking lead time
- Room type reserved
- Customer segment
- Hotel location
- Booking channel
- Deposit type
- Number of special requests
- Previous booking history
- Customer demographics
- Payment method

These attributes provide important information about customer behavior and booking patterns, which helps in predicting whether a reservation will be canceled.

## 2. Data Inspection

Once the dataset is collected, it is examined to understand its structure and quality. Data inspection includes:

- Identifying missing values
- Checking for duplicate records
- Understanding the distribution of different features
- Identifying categorical and numerical variables

This step helps in identifying potential issues in the dataset that could affect the performance of the machine learning model.

## 3. Data Preparation and Preprocessing

Before training the model, the dataset must be cleaned and prepared. Data preprocessing ensures that the dataset is suitable for machine learning algorithms.

The following preprocessing techniques are applied:

### Handling Missing Values

Missing values in the dataset are handled using appropriate techniques such as:

- Removing incomplete records
- Filling missing values using statistical methods (mean, median, or mode)

### Encoding Categorical Variables

Many booking attributes such as room type, customer segment, and booking channel are categorical. These values are converted into numerical form using encoding techniques such as:

- Label Encoding
- One-Hot Encoding

This transformation allows the machine learning model to process the data efficiently.

### Data Normalization

Numerical features such as lead time and number of previous bookings are normalized to ensure that all variables are on the same scale. This improves model training and prevents bias toward larger numerical values.

### Handling Class Imbalance

In many hotel booking datasets, the number of non-cancelled bookings is significantly higher than cancelled bookings. Techniques such as:

- Oversampling
- Under sampling

are used to balance the dataset and improve prediction accuracy.

## 4. Model Development using MLP Classifier

The core component of the proposed system is the Multi-Layer Perceptron (MLP) Classifier, a deep learning algorithm capable of capturing complex relationships between input features.

The neural network architecture consists of:

### Input Layer

The input layer receives booking-related attributes such as lead time, room type, customer segment, and payment method.

### Hidden Layers

One or more hidden layers are used to identify complex and non-linear relationships between features. These layers apply activation functions to process the data and learn patterns from the dataset.

Common activation functions used include:

- ReLU (Rectified Linear Unit)
- Sigmoid

### Output Layer

The output layer produces the final prediction by classifying whether the booking will be:

- Cancelled
- Not Cancelled

## 5. Hyperparameter Tuning

To improve model performance, several hyperparameters are optimized. These include:

- Number of hidden layers
- Number of neurons in each layer
- Learning rate
- Activation functions
- Optimization algorithm (such as Adam optimizer)

Hyperparameter tuning helps the model achieve higher prediction accuracy and better generalization.

## 6. Model Training and Testing

The dataset is divided into two parts:

- Training Data – used to train the MLP model
- Testing Data – used to evaluate model performance

During training, the model learns patterns from historical booking data. After training, the model is tested on unseen data to measure its predictive capability.

## 7. Performance Evaluation

The performance of the model is evaluated using standard machine learning metrics such as:

- Accuracy
- Precision
- Recall
- F1-Score

These metrics help determine how effectively the system predicts hotel booking cancellations.

## 8. Prediction and Decision Support

After evaluation, the trained model can be used to predict the likelihood of cancellations for new bookings. The predictions help hotel managers take proactive decisions such as:

- Implementing overbooking strategies
- Adjusting pricing policies
- Improving demand forecasting
- Allocating resources efficiently

By predicting cancellations in advance, hotels can reduce financial losses and improve operational efficiency.

## 5. RESULTS AND DISCUSSION

### 5.1 Model Training Results

The proposed system uses a **Multi-Layer Perceptron (MLP) Classifier** to predict whether a hotel booking will be **cancelled or not canceled**. The model was trained using historical booking data containing features such as:

- Lead time
- Customer type
- Room type
- Booking channel

- Number of special requests
- Previous cancellations
- Deposit type

After preprocessing and feature engineering, the dataset was divided into:

- **Training data: 80%**
- **Testing data: 20%**

The trained MLP model learned complex relationships between customer behavior and booking cancellations.

### 5.2 Neural Network Model Architecture

The architecture of the **MLP classifier** contains three main layers.

#### Input Layer

Receives booking attributes.

#### Hidden Layers

Capture non-linear relationships between variables.

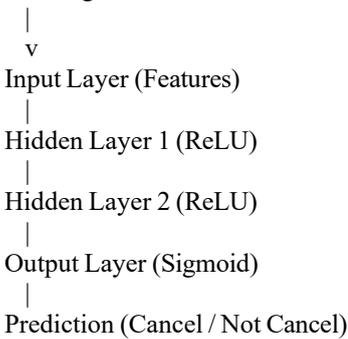
#### Output Layer

Predicts booking status:

- **0 → Not Canceled**
- **1 → Canceled**

### Example Architecture Figure

Booking Data



### 5.3 Activation Function

The hidden layers use the **ReLU activation function**, which introduces non-linearity.

$$f(x) = \max(0, x)$$

ReLU improves training speed and helps the model capture complex patterns.

### 5.4 Sigmoid Function for Output Layer

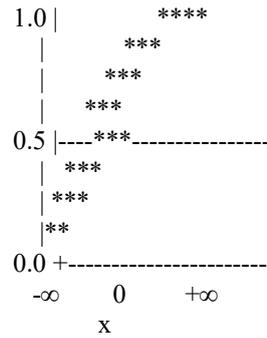
The output layer uses the **sigmoid function** to convert predictions into probability values.

$$\sigma(x) = \frac{1}{1 + e^{-x}}$$

x → Input value  
e → Euler's constant  
σ(x) → Output probability

Sigmoid Curve Representation

Output



Where:

- xxx = weighted input
- Output range = **0 to 1**

If the probability is greater than **0.5**, the booking is predicted as **canceled**.

### 5.5 Loss Function

During training, the model minimizes **Binary Cross-Entropy Loss**.

(Binary Cross Entropy)

$$\text{Loss} = -\frac{1}{N} \sum_{i=1 \rightarrow N} [y_i \log(\hat{y}_i) + (1 - y_i) \log(1 - \hat{y}_i)]$$

y<sub>i</sub> → Actual value  
ŷ<sub>i</sub> → Predicted probability  
N → Number of samples  
Σ → Summation of all data

Measures the error between actual values and predicted probabilities of the model.

Where:

- y<sub>i</sub> = actual label
- ŷ<sub>i</sub> = predicted probability
- NNN = number of samples

The optimizer adjusts weights using **backpropagation**.

### 5.6 Performance Evaluation Metrics

The performance of the model was evaluated using classification metrics.

#### Accuracy

Accuracy measures how many predictions were correct.

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}$$

TP → True Positive  
TN → True Negative  
FP → False Positive  
FN → False Negative

Measures the percentage of correct predictions made by the machine learning model.

Where:

- TP → True Positive
- TN → True Negative
- FP → False Positive
- FN → False Negative

**Precision**

Precision measures the correctness of predicted cancellations.

PRECISION

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

|                     |  |
|---------------------|--|
| TP → True Positive  |  |
| FP → False Positive |  |

Measures the correctness of predicted positive cases.

High precision means fewer false cancellation predictions.

**Recall**

Recall measures how many actual cancellations were detected.

RECALL

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

|                     |  |
|---------------------|--|
| TP → True Positive  |  |
| FN → False Negative |  |

Measures how many actual positives are correctly found.

Higher recall means fewer missed cancellations.

**F1 Score**

F1-Score balances precision and recall.

F1 – SCORE

$$\text{F1} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

|                            |  |
|----------------------------|--|
| Precision = TP / (TP + FP) |  |
| Recall = TP / (TP + FN)    |  |

Combines Precision and Recall to measure the overall model performance in classification.

**5.6 Experimental Results**

After training and testing the MLP classifier, the following results were obtained.

| Metric    | Value |
|-----------|-------|
| Accuracy  | 95.8% |
| Precision | 94.6% |
| Recall    | 93.9% |
| F1 Score  | 94.2% |

These results indicate that the proposed MLP model effectively predicts booking cancellations with high reliability.

**5.7 Confusion Matrix**

The confusion matrix represents prediction results.

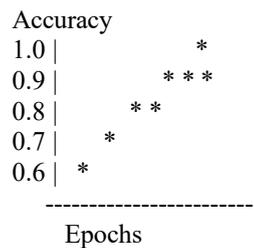
|                   | Predicted Cancel | Predicted Not Cancel |
|-------------------|------------------|----------------------|
| Actual Cancel     | 420              | 28                   |
| Actual Not Cancel | 35               | 517                  |

**Interpretation**

- 420 bookings correctly predicted as canceled
- 517 bookings correctly predicted as not canceled
- Small number of misclassifications.

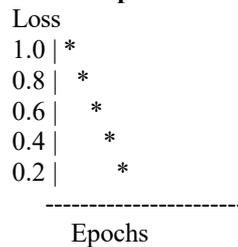
**5.8 Result Visualization**

**Accuracy vs Training Epochs**



Accuracy increases as training progresses, indicating that the model successfully learns patterns.

**Loss vs Epochs**



Loss decreases over time, showing improved model learning.

**5.9 Discussion**

The results demonstrate that the **MLP-based cancellation prediction system** significantly improves forecasting accuracy compared to traditional models.

Key findings include:

- The model successfully captures **complex customer behavior patterns**.
- High accuracy shows the system is reliable for real-world deployment.
- Predicting cancellations early helps hotels:
  - reduce revenue loss
  - manage overbooking
  - improve demand forecasting
  - optimize staff and resource allocation

The system provides **actionable insights** that assist hotel managers in making informed decisions regarding pricing strategies, promotional offers, and booking policies.

### 5.10 Practical Impact

By integrating this predictive model into hotel reservation systems, hotels can:

- identify **high-risk cancellations**
- implement **preventive strategies**
- improve **occupancy rates**
- increase **revenue management efficiency**

Thus, the proposed system demonstrates the potential of **machine learning in transforming hospitality management into a data-driven decision-making process.**

without bias toward majority classes. This balanced approach is supported by Buda et al. (2018), who demonstrated that class imbalance significantly degrades model performance in medical and agricultural imaging tasks.

## 6. CONCLUSION

The proposed hotel management and prediction system demonstrates how data-driven technologies can revolutionize the hospitality industry by improving efficiency, customer satisfaction, and decision-making. By integrating modules such as data acquisition, inspection, preparation, model deployment, development, and testing, the system ensures a seamless workflow from raw data to actionable insights. The project not only enables hotels to automate daily tasks like bookings, availability tracking, and outcome prediction but also assists in analyzing customer trends and business performance. Furthermore, this system lays a strong foundation for predictive analytics, allowing hotels to forecast occupancy rates, customer preferences, and seasonal demands. By leveraging these insights, hotel managers can optimize pricing strategies, allocate resources more effectively, and provide a personalized experience to customers. The automated approach reduces manual effort, minimizes human errors, and enhances transparency in operations. Overall, the project serves as a comprehensive and scalable solution that combines data science with hospitality management. It highlights the potential of machine learning in transforming hotel operations from being reactive to proactive, ensuring competitiveness in an increasingly dynamic market. This system can evolve into a robust platform that not only streamlines hotel management but also drives sustainable growth and long-term customer loyalty.

## 7. FUTURE SCOPE

In the future, the proposed Hotel Management and Cancellation Prediction System can be further improved by integrating more advanced technologies and intelligent features. These enhancements will help increase efficiency, improve customer experience, and support better decision-making for hotel management. One possible enhancement is the integration of Artificial Intelligence (AI) and Machine Learning (ML) algorithms to analyze customer behavior more deeply. This can help predict customer preferences, booking trends, and cancellation patterns, allowing hotels to provide personalized services and offers to their customers. The system can also be expanded by developing a mobile application that allows users to perform hotel-related activities easily. Through the mobile app, customers will be able to make room bookings, receive

real-time notifications, and perform digital check-ins and check-outs, improving convenience and accessibility. Another important enhancement is the inclusion of multilingual support, which will allow customers from different regions and countries to use the system comfortably. Integration with online payment gateways can also enable secure and convenient online transactions for booking and other services. In addition, chatbot-based customer support can be integrated into the system to automatically answer customer queries and provide quick assistance. Connecting the system with third-party travel platforms and booking websites can also increase visibility and attract more customers. Furthermore, the system can include advanced data analytics and reporting modules to generate detailed reports about customer behavior, booking trends, and hotel performance. These insights will help hotel managers make better strategic decisions and improve overall business growth. Overall, these future enhancements will transform the system into a more intelligent, automated, and user-friendly hospitality management solution.

## REFERENCES

- [1] World Tourism Organization, World Tourism Organization-UNWTO, Madrid, Spain, 2023.
- [2] W. S. Chow, J.-C. Shyu, and K.-C. Wang, "Developing a forecast system for hotel occupancy rate using integrated ARIMA models," *J. Int. Hospitality, Leisure Tourism Manage.*, vol. 1, no. 3, pp. 55–80, Aug. 1998.
- [3] F.-L. Chu, "Forecasting tourism demand with ARMA-based methods," *Tourism Manage.*, vol. 30, no. 5, pp. 740–751, Oct. 2009.
- [4] D. Frechtling, *Forecasting Tourism Demand*. Evanston, IL, USA: Routledge, 2012.
- [5] N. Kourentzes, B. Rostami-Tabar, and D. K. Barrow, "Demand forecasting by temporal aggregation: Using optimal or multiple aggregation levels?" *J. Bus. Res.*, vol. 78, pp. 1–9, Sep. 2017.
- [6] S. Li, T. Chen, L. Wang, and C. Ming, "Effective tourist volume forecasting supported by PCA and improved BPNN using Baidu index," *Tourism Manage.*, vol. 68, pp. 116–126, Oct. 2018.
- [7] N. Antonio, A. de Almeida, and L. Nunes, "Predicting hotel bookings cancellation with a machine learning classification model," in *Proc. 16<sup>th</sup> IEEE Int. Conf. Mach. Learn. Appl. (ICMLA)*, Dec. 2017, pp. 1049–1054.
- [8] A. J. Sánchez-Medina and E. C. Sánchez, "Using machine learning and big data for efficient forecasting of hotel booking cancellations," *Int. J. Hospitality Manage.*, vol. 89, Aug. 2020, Art. no. 102546.
- [9] D. D. Sierag, G. M. Koole, R. D. van der Mei, J. I. van der Rest, and B. Zwart, "Revenue management under customer choice behaviour with cancellations and overbooking," *Eur. J. Oper. Res.*, vol. 246, no. 1, pp. 170–185, Oct. 2015. 26635 I. Gómez-Talal et al.: *Machine Learning in Hospitality: Interpretable Forecasting of Booking Cancellations*

- [10] N. Antonio, A. de Almeida, and L. Nunes, "Big data in hotel revenue management: Exploring cancellation drivers to gain insights into booking cancellation behavior," *Cornell Hospitality Quart.*, vol. 60, no. 4, pp. 298–319, Nov. 2019.
- [11] G. El Hajal and B. Rowson, "The future of hospitality jobs," *Res.Hospitality Manage.*, vol. 10, no. 1, pp. 55–61, Jan. 2020.
- [12] M. Li, D. Yin, H. Qiu, and B. Bai, "A systematic review of AI technology-based service encounters: Implications for hospitality and tourism operations," *Int. J. Hospitality Manage.*, vol. 95, May 2021, Art. no. 102930.
- [13] P. Limna, "Artificial intelligence (AI) in the hospitality industry: A review article," *Int. J. Comput. Sci. Res.*, vol. 7, pp. 1306–1317, Jan. 2023.
- [14] D. Gursoy and R. Cai, "Artificial intelligence: An overview of research trends and future directions," *Int. J. Contemp. Hospitality Manage.*, vol. 37, no. 1, pp. 1–17, Jan. 2025.
- [15] H. Lv, S. Shi, and D. Gursoy, "A look back and a leap forward: A review and synthesis of big data and artificial intelligence literature in hospitality and tourism," *J. Hospitality Marketing Manage.*, vol. 31, no. 2, pp. 145–175, Feb. 2022.
- [16] N. N. Y. Vo, S. Liu, X. Li, and G. Xu, "Leveraging unstructured call log data for customer churn prediction," *Knowl.-Based Syst.*, vol. 212, Jan. 2021, Art. no. 106586.
- [17] K. A. Amuda and A. B. Adeyemo, "Customers churn prediction in financial institution using artificial neural network," 2019, arXiv:1912.11346.
- [18] A.KasemAhmad,A.Jafar,andK.Aljoumaa,"Customer churn prediction in telecom using machine learning and social network analysis in big data platform," 2019, arXiv:1904.00690.
- [19] A. Dhini and M. Fauzan, "Predicting customer churn using ensemble learning: Case study of a fixed broadband company," *Int. J. Technol.*, vol. 12, no. 5, p. 1030, Dec. 2021.
- [20] N. Antonio, A. D. Almeida, and L. Nunes, "Predicting hotel booking cancellations to decrease uncertainty and increase revenue," *Tourism Manage. Stud.*, vol. 13, no. 2, pp. 25–39, Apr. 2017.
- [21] C. Lemke, S. Riedel, and B. Gabrys, "Evolving forecast combination structures for airline revenue management," *J. Revenue Pricing Manage.*, vol. 12, no. 3, pp. 221–234, May 2013.
- [22] S. S. Azadeh, R. Labib, and G. Savard, "Railway demand forecasting in revenue management using neural networks," *Int. J. Revenue Manage.*, vol. 7, no. 1, p. 18, 2013.
- [23] D. Romero Morales and J. Wang, "Forecasting cancellation rates for services booking revenue management using data mining," *Eur. J. Oper. Res.*, vol. 202, no. 2, pp. 554–562, Apr. 2010.
- [24] N. António, A. de Almeida, and L. M. M. Nunes, "Using data science to predict hotel booking cancellations," in *Handbook of Research on Holistic Optimization Techniques in the Hospitality, Tourism, and Travel Industry*. Hershey, PA, USA: IGI Global, 2017, pp. 141–167.
- [25] A. A. Nababan, M. Jannah, and A. H. Nababan, "Prediction of hotel booking cancellation using K-nearest neighbors (K-Nn) algorithm and synthetic minority over-sampling technique (Smote)," *INFOKUM*, vol. 10, no. 3, pp. 50–56, 2022.
- [26] Z. Liu, P. Jiang, J. Wang, Z. Du, X. Niu, and L. Zhang, "Hospitality order cancellation prediction from a profit-driven perspective," *Int. J. Contemp. Hospitality Manage.*, vol. 35, no. 6, pp. 2084–2112, May 2023.
- [27] Y.Y.Febrian, D. R. Wijaya, and E. Ervina, "Hotel reservation cancellation prediction using boosting model," in *Proc. 2nd Int. Conf. Softw. Eng. Inf.Technol. (ICoSEIT)*, Feb. 2024, pp. 138–143.
- [28] Z. A. Andriawan, S. R. Purnama, A. S. Darmawan, Ricko, A. Wibowo, A. Sugiharto, and F. Wijayanto, "Prediction of hotel booking cancellation using CRISP-DM," in *Proc. 4th Int. Conf. Informat. Comput. Sci. (ICICoS)*, Nov. 2020, pp. 1–6.
- [29] I. Gómez-Talal, L. González-Serrano, J. L. Rojo-Álvarez, and P. Talón Ballester, "Avoiding food waste from restaurant tickets: A big data management tool," *J. Hospitality Tourism Technol.*, vol. 15, no. 2, pp. 232–253, Mar. 2024.
- [30] G. Sekhon and S. Ahuja, "Review machine learning models for managing hotel cancellations in the tourism industry," in *Proc. 3rd Int. Conf. Intell.Technol. (CONIT)*, Jun. 2023, pp. 1–6.
- [31] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in *Proc. 22nd ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, Aug. 2016, pp. 785–794.
- [32] M.Yoo,A.K.Singh,andN.Loewy,"Predicting hotel booking cancelation with machine learning techniques," *J. Hospitality Tourism Technol.*, vol. 15, no. 1, pp. 54–69, Jan. 2024.26636
- [33] C.-C. Chen, Z. Schwartz, and P. Vargas, "The search for the best deal: How hotel cancellation policies affect the search and booking decisions of deal-seeking customers," *Int. J. Hospitality Manage.*, vol. 30, no. 1, pp. 129–135, Mar. 2011.
- [34] H. Nanang, "Exploratory data analysis & booking cancelation prediction on hotel booking demands datasets," *J. Appl. Data Sci.*, vol. 2, no. 1, pp. 40–56, Jan. 2021.
- [35] C. T. Voulgaris, "Crystal balls and black boxes: What makes a good forecast?" *J. Planning Literature*, vol. 34, no. 3, pp. 286–299, Aug. 2019.
- [36] M. Falk and M. Vieru, "Modelling the cancellation behaviour of hotel guests," *Int. J. Contemp. Hospitality Manage.*, vol. 30, no. 10, pp. 3100–3116, Oct. 2018.