Upswing Hotel booking executive summery

Executive Summary

This report analyzes hotel cancellations using predictive modeling techniques. By employing Logistic Regression and Random Forest models, significant factors influencing cancellations were identified. The models achieved high accuracy rates (98-99%), indicating their effectiveness in predicting cancellations. Actionable insights and recommendations are provided for hotel management to minimize cancellations and optimize revenue.

Introduction

Hotel cancellations significantly impact revenue and operational efficiency. Understanding the factors leading to cancellations enables hotel management to implement strategies that minimize their occurrence, thereby improving overall business performance.

Data Description

The dataset consists of booking details from a resort hotel, including variables such as lead time, meal type, customer segment, and cancellation status. Key features include lead-time, adults, children, country, reservation status, and average daily rate (ADR). This data is essential for building a predictive model for cancellations.

Methodology

The analysis involved the following steps: selecting relevant features, splitting the dataset into training and test sets, and training two models: Logistic Regression and Random Forest. Model performance was evaluated using metrics such as accuracy, precision, recall, and F1-score.

Data Preprocessing

Data preprocessing steps included handling missing values, encoding categorical variables, and normalizing numerical features. This step is critical to ensure that the data is suitable for modelling and can lead to better predictive performance.

Model Training and Evaluation

The training process involved fitting the Logistic Regression and Random Forest models on the training data. The evaluation metrics indicated that both models performed well, with accuracy rates around 98-99%. Additional metrics such as precision, recall, and F1-score were calculated to provide a comprehensive view of model performance.

Analysis and Results

The analysis revealed key insights, including the strong correlation between lead time and cancellation likelihood. Furthermore, customer segmentation analysis identified specific groups with higher cancellation rates, informing targeted marketing and pricing strategies.

Operational Insights and Recommendations

Based on the analysis, it is recommended to implement flexible pricing strategies, target high-risk customer segments with tailored offers, enhance customer experience through easy modifications, and use real-time cancellation tracking for informed decision-making.

Conclusion

This report emphasizes the importance of predictive analysis in managing hotel cancellations. Future work could focus on refining models with additional data and exploring advanced machine learning techniques to enhance predictive accuracy.