Project: INN Hotels

Analyzed the data of INN Hotels to find which factors have a high influence on booking cancellations, build a predictive model that can predict which booking is going to be canceled in advance, and help in formulating profitable policies for cancellations and refunds.

Covered Skills and Tools

- Initiated the project with a detailed Exploratory Data Analysis (EDA) to define the problem of booking cancellations at INN Hotels. Utilized univariate and bivariate analyses to unearth patterns and insights from the booking data, employing various visualizations to highlight key observations about individual variables and their interactions.
- Executed comprehensive data preprocessing to prepare the dataset for modeling. This included treating missing values, detecting and handling outliers, and performing feature engineering to enhance the predictive power of the models. Split the data into training and testing sets to ensure a robust evaluation of model performance.
- Developed a Logistic Regression model using the stats-models library, thoroughly analyzing model statistics to assess fit and performance. Checked for adherence to Logistic Regression assumptions and interpreted the influence of different variables on booking cancellations through the coefficients obtained.
- Enhanced the Logistic Regression model's performance by adjusting the classification threshold, aiming to balance precision and recall effectively. Provided a detailed analysis of the impact of this adjustment on model performance, ensuring an optimal threshold for predicting cancellations.
- Constructed a Decision Tree model to explore an alternative predictive approach. Evaluated the initial model's performance and embarked on model optimization through pruning, reducing complexity while maintaining accuracy. Analyzed feature importance and extracted decision rules to gain deeper insights into factors influencing cancellations.
- Conducted a comparative analysis of the Logistic Regression and Decision Tree models, highlighting their strengths and limitations in the context of predicting booking cancellations. This comparison facilitated a comprehensive understanding of model suitability for different aspects of the cancellation prediction task.
- Derived actionable insights and formulated strategic recommendations for INN Hotels to mitigate the impact of booking cancellations. Emphasized the importance of understanding customer behavior patterns and implementing data-driven policies for cancellations and refunds to enhance revenue management.
- Maintained high standards in the presentation of the project, ensuring the notebook was well-structured with a clear and logical flow. Emphasized visual appeal through crisp, well-

commented code and compelling visualizations. Concluded with concise and actionable business recommendations, underlining the practical value and implications of the predictive modeling efforts.

- Other Covered Skills: Exploratory Data Analysis, Data Preprocessing, Logistic Regression, Multi-collinearity, AUC-ROC Curve, Decision Tree, Pruning.