**INN-Hotels**

Machine Learning based solution that can help in predicting which booking is likely to be canceled for INN Hotels group.

EDA is done. Outliers were treated. Encoded categorical features. Identified F1 as evaluation metric.

Built Logistic regression model. Checked for multicollinearity. Dropped features with high p-values to ensure there is no multicollinearity. Built model with selected features. Evaluated the model. Improved the F1-score by changing model threshold. Used precision-recall curve to find better threshold.

Built Decision Tree model. Evaluated the metrics. Identified important features. Tuned the model with hyperparameters. Evaluated metrics. Pruned the tree with ccp alpha. Evaluated metrics. Compared all the 3 models. Selected the best model. Identified important features. Made recommendations based on that.