Background Study

Classification Algorithms like the Random Forest Classifier are designed to assign predefined labels to input data based on patterns identified during the training phase. The goal of a classifier is to learn a mapping between input features and the corresponding output labels so that it can make accurate predictions on new, unseen data. Random Forest Classifier is an ensemble learning method, built on the foundation of decision trees. Random Forest creates multiple subsets of the training data, where each subset is used to train a decision tree. Features like number of people, arrival date, special services, car parking, etc. are essential in determining the outcome of a hotel reservation. The algorithm classifies the result in two sets: likely to cancel, and not likely to cancel. In order to test the performance of the model, multiple performance metrics like Accuracy, Precision, Recall, F1-Score, and the AUC-ROC Curve will be measured.