# Abstract:

**Booking cancellations are undoubtedly one the biggest headaches of any revenue manager or hotel manager nowadays. So our goal for this Classification Model is to predict the individuals that are going to cancel their hotel booking. This model will allow hotels to predict if a new booking will be canceled or not, manage their business accordingly, and increase their revenue.**

**Design:**

**We used several machine learning algorithms such as logistic regression, k-nearest neighbors, decision tree, and random forest classifiers**

**Data:**

**Data we used in this article is publicly available at Kaggle, this data set contains a single file which compares various booking information between two hotels: a city hotel and a resort hotel. The dataset contains 119364 points with 32 features for each. A few feature highlights include lead time, deposit type and is canceled**

**Algorithms:**

**Data Preparation :**

**• Handle missing values**

**• Encode categorical data**

**• Fix the data types**

**• Handling features**

**• Correlation**

**Models**

**• Logistic Regression**

**• Decision Tree Model**

**•Random Forest Classifier**

**•k-nearest neighbors**

**The model performance is measured using accuracy, f1, recall, precision, and the confusion matrix..Decision Tree is the best model.**

**Tools:**

### **Technologies:**

* **Python**
* **Jupyter Notebook**

### **Libraries:**

* **Pandas**
* **Numpy**
* **Matplotlib**
* **Seaborn**
* **Sklearn**

**Communication:**

**our project will be included on our GitHub page. Thank YOU SADIA**