Hotel Booking Prediction using Machine Learning Algorithms

1. Introduction:

The goal of this project is to create a machine learning model that, using the dataset, can reliably predict hotel booking. The dataset includes booking details for a resort hotel and a city hotel. These details include the booking date, the number of guests, the length of the stay, and the availability of parking. To create prediction models, the research will make use of three well-known machine learning algorithms: CatBoost, XGBoost, and Random Forest.

2. Problem Statement:

To optimize resource allocation, personnel, and revenue management, the hotel sector primarily depends on accurate demand forecasts. We want to help hotel management make educated decisions and improve operational efficiency by creating a strong predictive model. The likelihood that a reservation would be canceled will be predicted by the model.

3. Methodology:

The project will follow the following steps:

3.1 Data Preprocessing:

- Use exploratory data analysis (EDA) to discover new information about the dataset.
- Handle any data inconsistencies, outliers, and missing values.
- Create training and testing sets from the dataset.

3.2 Model Development:

- On the preprocessed dataset, train the three machine learning algorithms CatBoost, XGBoost, and Random Forest.
- To enhance model performance, undertake hyperparameter tuning for each algorithm using strategies like grid search or random search.
- Evaluate the model's performance using appropriate evaluation metrics including accuracy, precision, recall, and F1-score.
- Evaluate the three algorithms' performance and choose the model that performs the best.

4. Expected Deliverables:
- A thorough report describing the project's methodology, including the stages involved in data preprocessing, feature engineering methods, model building, evaluation outcomes, and key learnings.
- CatBoost, XGBoost, and Random Forest are three trained machine learning models that can forecast hotel booking.
- Presentation slides that list the project's key findings.

5. Timeline:
- Week 1: Data exploration and preprocessing, Feature engineering and model development.
- Week 2: Hyperparameter tuning and model evaluation. completing the presentation and project report.

Note: All project members will be involved in all the project steps.

6. Required Resources:
- The provided hotel booking dataset.
- Python programming language with libraries such as pandas, scikit-learn, CatBoost, XGBoost, RandomForest and matplotlib for data visualization.
- An appropriate development environment (such as PyCharm or Jupyter Notebook).

7. Conclusion:
This study intends to create precise predictive models for hotel booking by utilizing machine learning techniques including CatBoost, XGBoost, and Random Forest. The project's results will support hotel management in enhancing revenue management techniques, customer service.

Group Info.

Group - 1		
Name	**ID**	**Mail Ids**
Ketha Tirumuru	11597873	kethatirumuru@my.unt.edu
Chandana Polakonda	11645295	chandanapolakonda@my.unt.edu
Arun Thotakuri	11600012	arunthotakuri@my.unt.edu
Numitha Devi Oguri	11647630	numithadevioguri@my.unt.edu
Lalitha Nali	11649335	lalithanali@my.unt.edu