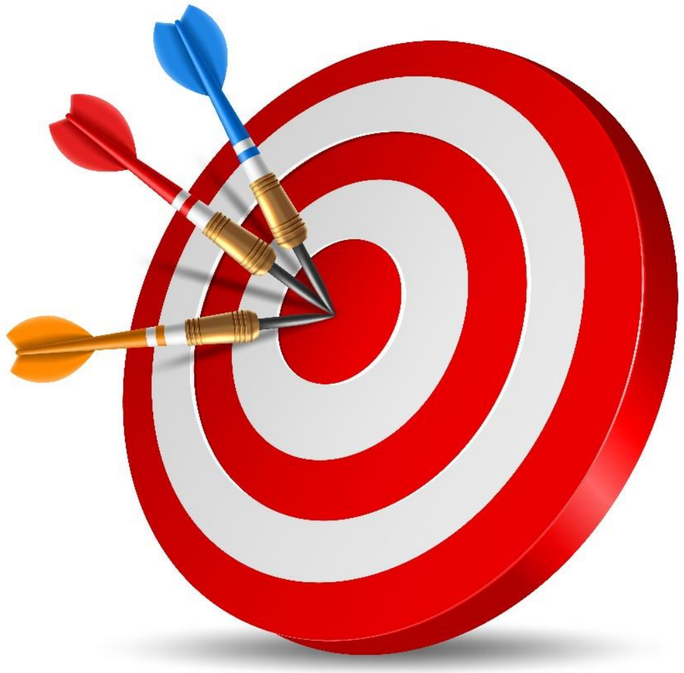
After Module 2

Assignment 1: Build a Predictive Model to Forecast Hotel Occupancy

Objective:

To create a predictive model that forecasts hotel occupancy using publicly available data and AI tools like Python or DataRobot.



Instructions

Dataset Recommendation:

Use the Kaggle dataset on hotel bookings. This dataset includes information like booking dates, room types, lead time, and more. (You can also synthetically generate the data) or use any other dataset which has granular details of revenue.

Download the dataset and review the documentation for variables and data structure.

Steps to Complete the Assignment:

Step 1: Data Preparation

- Load the dataset into Python using libraries like Pandas.
- Clean the data by removing duplicates, handling missing values, and encoding categorical variables.
- Explore the data to identify key features that may influence hotel occupancy, such as booking lead time, market segment, and seasonality

Step 2: Choose a Predictive Model

Select a machine learning model, such as linear regression, random forests, or an ensemble model like Gradient Boosting Machines (GBM).

Use a library like Scikit-learn to train the model.

Step 3: Train and Test

Split the data into training and testing sets.

Train the model on the training set and evaluate its accuracy using the testing set. Use metrics like Mean Absolute Error (MAE) or R-squared.

Step 4: Visualize Predictions

Create visualizations (e.g., line graphs) comparing actual and predicted occupancy rates.

Step 5: Submit Findings

Share your model's performance metrics, visualizations, and a summary of your findings.