**Course Work Assessment 1**

## **Module Title:** Computing for AI

## **Module Code:** CMP6221

## **School:** School of Computing and Digital Technology

## **Assessment Title:** Portfolio of Secured Database-enabled programs

## **Module Coordinator:** Salameh Abu Rmeileh

## **Student Name:** Chaithanya Vamshi Sai

## **Student Id:** 211527297

## **Assessment 1 Summary:**

**Description:** Assessment 1 is a portfolio of python solutions for the lab exercise and the case study. Portfolio should include all the labs solutions and your solution to the case study as Python programs.

**Case study:** Develop a Decision Tree Classifier Algorithm in Python using to analyze a dataset.

**Case Study Implementation Summary of Assessment 1:**

1. Developed **Decision Tree Classifier using Machine Learning with Python** to solve real world problem from a Kaggle dataset.
2. Implemented **Machine Learning Project Life Cycle** by using **CRISP - DM** methodology to solve business problem.

Below are steps implemented in building Decision Tree Classifier using Machine Learning with Python.

## **Machine Learning Project Life Cycle:**

1. Understanding the Business Problem
2. Data Collection and Understanding
3. Data Exploration
4. Data Preparation
5. Modeling
6. Model Deployment

**Project Case Study:**

# **Predicting Hotel Booking Cancellations Using Machine Learning with Python**

**Guidelines followed in implementation of Case Study:**

1. **Understanding Business Problem**

* Have you ever wondered what if there was a way, we could predict which guests are likely to Cancel the Hotel Booking? That would be great right?
* The Goal of this project is to Predict the Guests who are likely to Cancel the Hotel Booking using Machine Learning with Python. Therefore, predicting reservations which might get canceled and preventing these cancellations will create a surplus revenue, better forecasts and reduce uncertainty in business management decisions.

1. **Data Collection and Understanding**

* After understanding the Business problem, next step is to collect and understand the data. The dataset is collected from the Kaggle.

<https://www.kaggle.com/jessemostipak/hotel-booking-demand>

1. **Data Exploration**

* Importing Libraries
* Store the data into MongoDB Database
* Load (Fetching) data from MongoDB Database
* Exploratory Data Analysis (EDA) on all Features
* Data Visualisation on all Important Features

1. **Data Preparation**

* Handling Missing Values
* Exploring Numerical and Categorical Features
* Feature Engineering (Encoding Categorical Features)
* Feature Selection (Correlation Heat Map)

1. **Modeling**

* Model Training
* Visualizing Decision Trees
* Feature Importance
* Model Evaluation
* Optimizing Model Performance

1. **Model Deployment**

* In this project, I have deployed Machine Learning Model on Heroku Cloud. Once we have deployed it on the cloud, we have successfully built a Data Science product that can be used by the Business User/ Customer.
* Machine Learning Web App is deployed on Heroku Cloud using Python and Streamlit.
* Link to the ML Web app: <https://cancelation-predictor.herokuapp.com/>

**Requirements Covered in this Project for Case Study Assessment 1:**

1. Understand the Business Problem to solve real world problem ✅
2. Collected the real-world case study dataset from Kaggle ✅
3. Mongo DB database is Username and Password protected ✅
4. Data is stored in MongoDB Atlas Cloud ✅
5. Retrieved (Fetch) data from MongoDB Atlas Cloud ✅
6. Exploratory Data Analysis (EDA) on all Features ✅
7. Data Visualisation on all Important Features ✅
8. Performing Feature Engineering on dataset ✅
9. Implementing Decision Tree Algorithm on the given Dataset ✅
10. Evaluating the Decision Tree Classifier Model ✅
11. Visualize Decision Tree Graph ✅
12. Optimizing Decision Tree Performance by Hyper Parameter Tuning ✅
13. Designed ML Web Application (GUI) for the python program to display data results and interact with front end user ✅
14. Deployed the ML Web Application on the cloud to make it scalable and secured ✅
15. Program is Username and Password protected ✅
16. Program is in OOP structure ✅
17. Program is functional ✅
18. Clear and Consistent naming convention for functions, variables and classes ✅
19. Code is Commented for understanding of theoretical concepts and implementation of code ✅
20. Program has functionality to run on a fresh installation of Python and MongoDB ✅

**GitHub Link for the Python Code for Assessment 1 Case Study:**

<https://github.com/ChaithanyaVamshi/Hotel_Cancelation_Predictor_ML_WebAPP>

**ML Web App link deployed on Heroku Cloud:**

https://cancelation-predictor.herokuapp.com/