Question – 1:

For the given RealEstate csv, write a python program satisfying following tasks to demonstrate application of machine learning through
multiple linear regression as follows –
Given: -
Dataset RealEstate.csv
ML Library to be used scikit-learn
Dependent variable 'Y house price of unit area'
Independent variables 'X1 transaction date', 'X2 house age', 'X3 distance to the nearest MRT station', 'X4 number of convenience stores', 'X5 latitude' and 'X6 longitude'
1. Import required libraries.
2. Load RealEstate dataset, create a dataframe and check datatypes of its attributes using appropriate method.
3. Remove 'No' column from the dataframe.
4. Check for any null values in features using appropriate method.
5. Create feature variables x and y as given above.
6. Create training and testing sets of feature variables with 70% of data for training and with random state of 110.
7. Create and fit regression model using appropriate method.
8. Use testing set created in step 6 to find and print the prediction of the outcome.
9. Find and print coefficient and mean squared error of the regression model.

Question – 2:

The objective is to perform classification on the Iris dataset using the k-Nearest Neighbors (kNN) algorithm. The Iris dataset contains measurements of various iris flowers, including features such as sepal length, sepal width, petal length, and petal width, along with the corresponding species label. The problem involves two main tasks:

- Build a kNN classification model to predict the species of iris flowers based on their feature measurements.
- Train the model on a portion of the dataset and evaluate its performance on another portion to assess its accuracy.
- Experiment with different values of k and choose the optimal value that maximizes the model's performance.
- Use appropriate evaluation confusion matrix to evaluate the model's performance. Also calcualte accuracy, sensitivity and specificity.
- Use iris.csv file for dataset.

Question – 3:

The task involves building a Decision Tree classifier to predict whether to play tennis based on weather conditions. The dataset used for this task is the PlayTennis dataset, which contains information about various weather attributes such as outlook, temperature, humidity, and wind, along with the corresponding decision to play tennis or not. Use PlayTennis.csv for dataset.