**OJT PRACTICAL EXAMINATION**

**Name: Sreethu Mohan**

**Institute: NSTI (W) TVM**

**Aim**:

Using a given dataset, apply various machine learning techniques to classify and predict outcomes. Evaluate the performance of your models using different statistical methods, confusion matrix, and cross-validation.

**List of Hardware and Software Requirements:**

1. Windows 10 or 11
2. Jupyter Notebook/Jupyter Lab
3. V S code
4. Python

**PROCEDURE:**

**Step 1:** Open Anaconda navigator

**Step 2:** Launch Jupyter Lab

**Step 3:** Create a new Python file on the folder you want to save it.

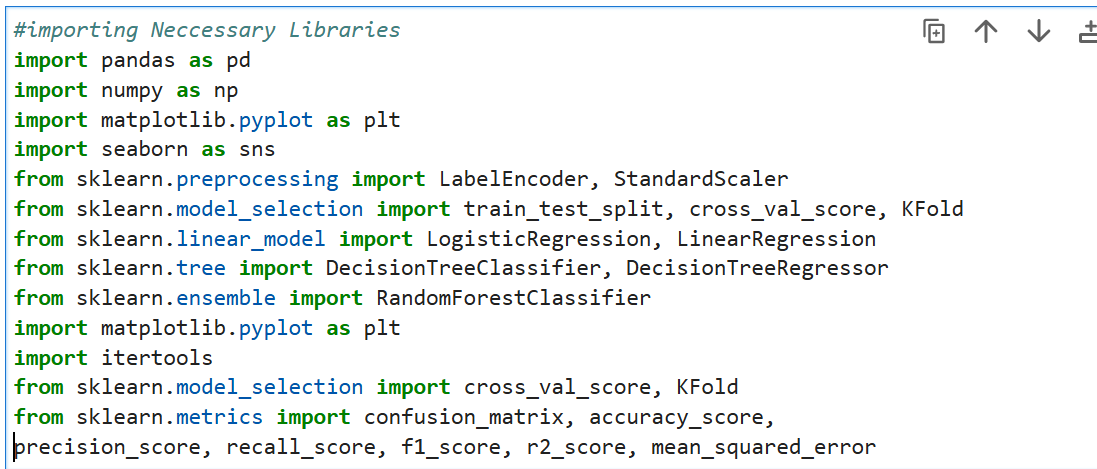
**Step 4:** Rename the file and type the code to execute the program in the Jupyter Lab tab

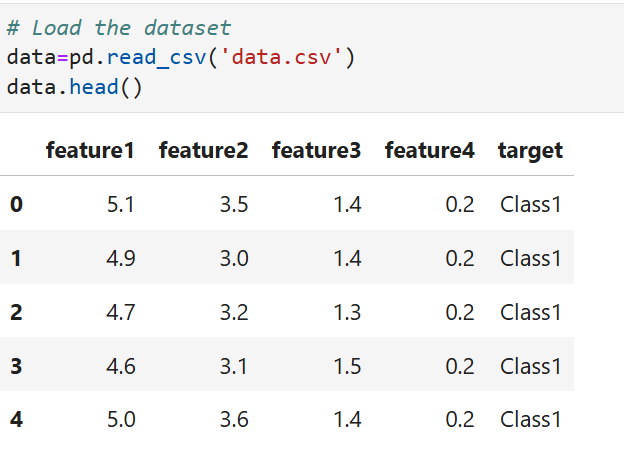
**Step 5:** Save and run the code

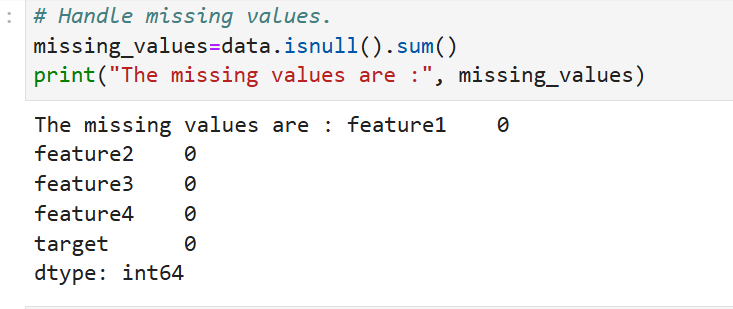
**Task1:**

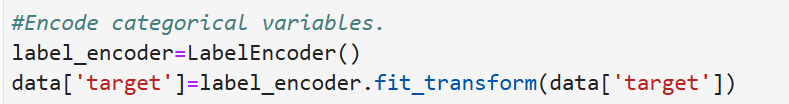
1. Data Preprocessing:
2. Load the dataset.
3. Handle missing values.
4. Encode categorical variables.
5. Scale/normalize the features

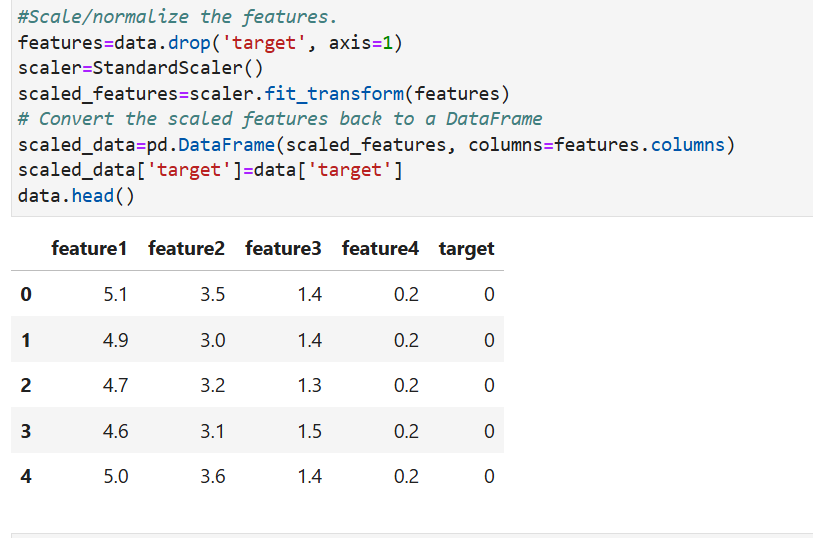
**Code:**

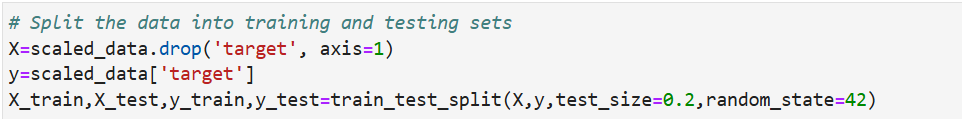
****

****

****

****

****

****

**Result:**

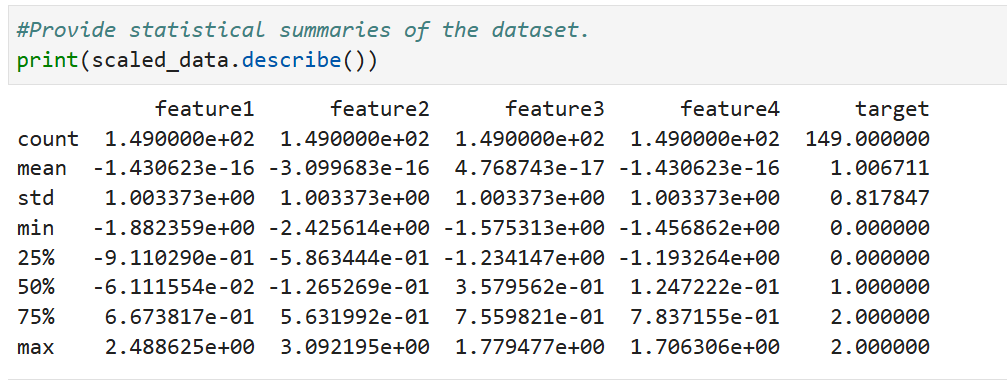
All the given task run successfully

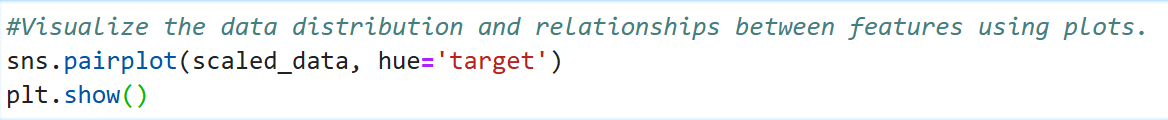
**Task2:**

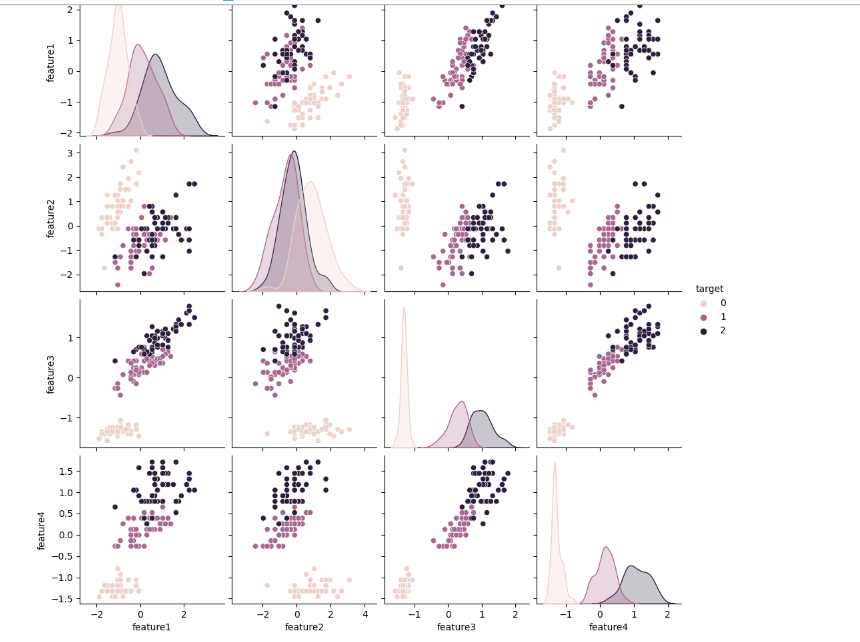
Exploratory Data Analysis (EDA):

* Provide statistical summaries of the dataset.
* Visualize the data distribution and relationships between features using plots.

**Code:**

****

****

****

**Result:**

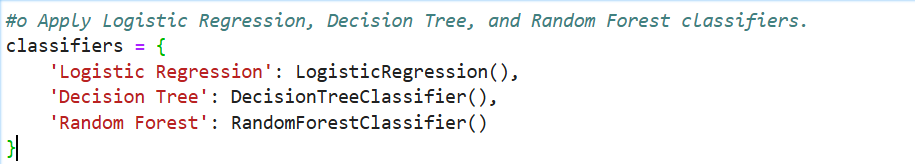
All the given task run successfully

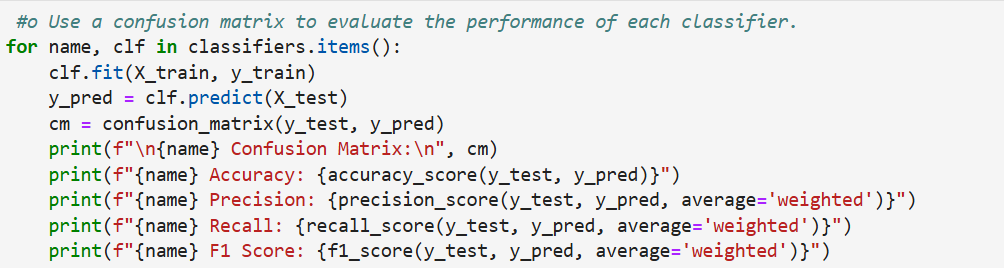
**Task3:**

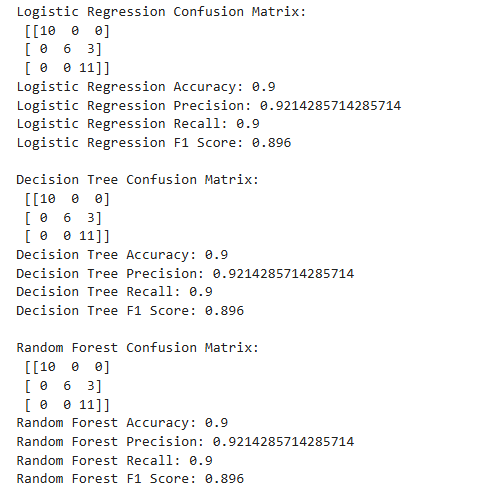
Classification:

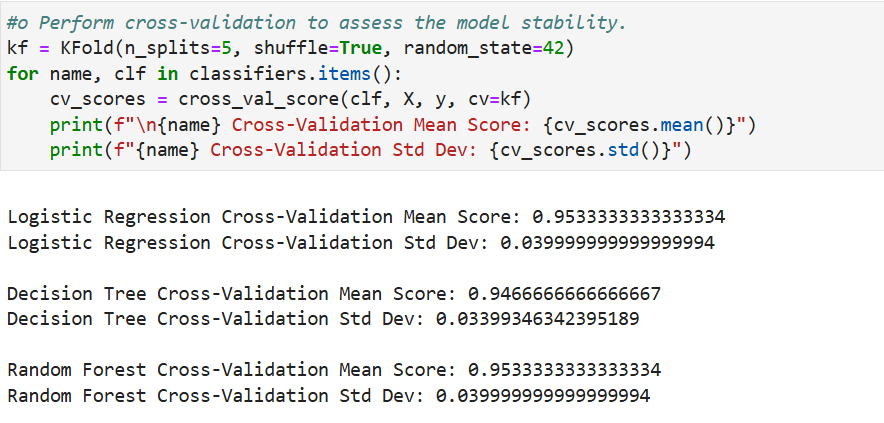
* Apply Logistic Regression, Decision Tree, and Random Forest classifiers.
* Use a confusion matrix to evaluate the performance of each classifier.
* Perform cross-validation to assess the model stability.

**Code**

****

****

****

****

**Result:**

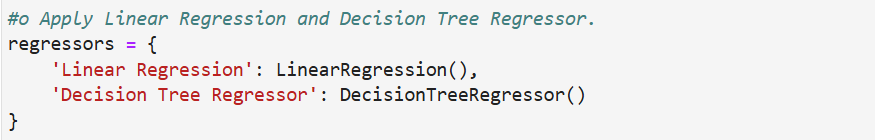
All the given task run successfully

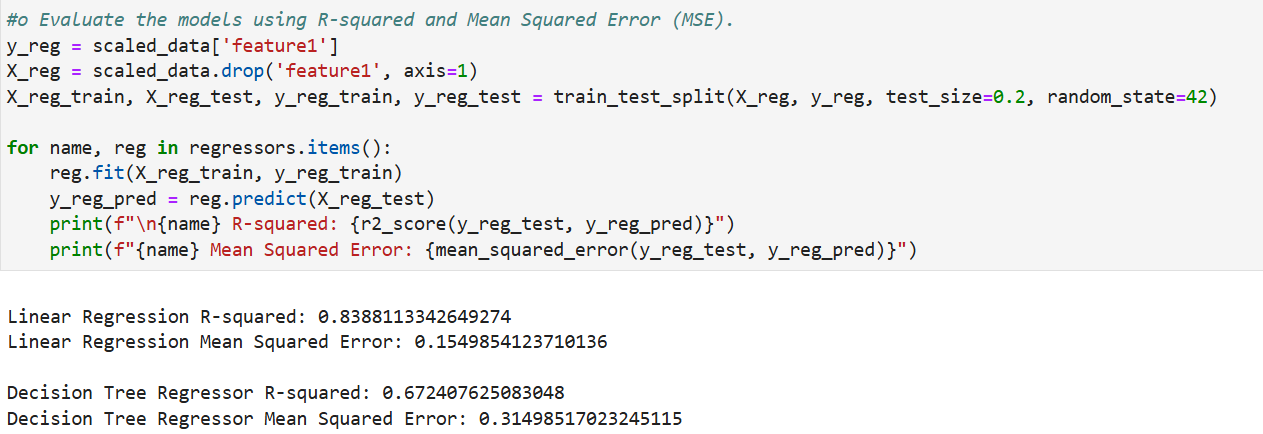
**Task4:**

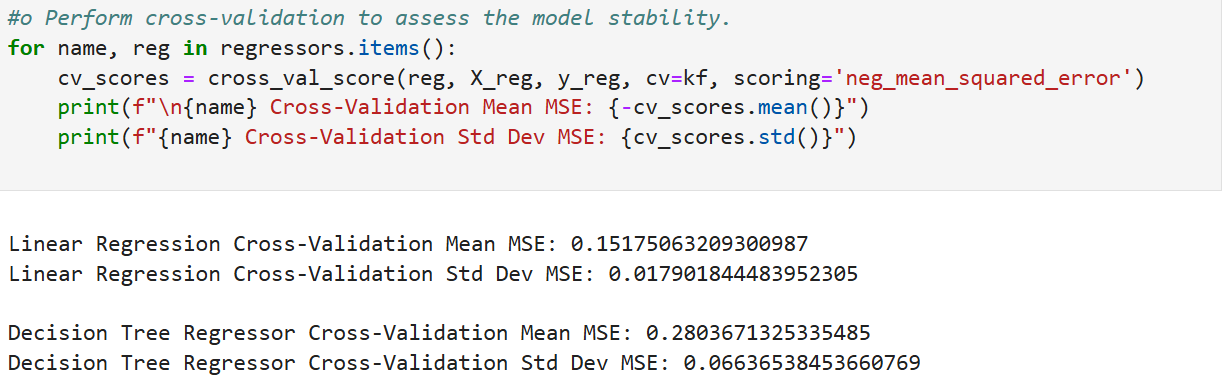
Regression:

* Apply Linear Regression and Decision Tree Regressor.
* Evaluate the models using R-squared and Mean Squared Error (MSE).
* Perform cross-validation to assess the model stability.

**Code:**

****

****

****

**Result:**

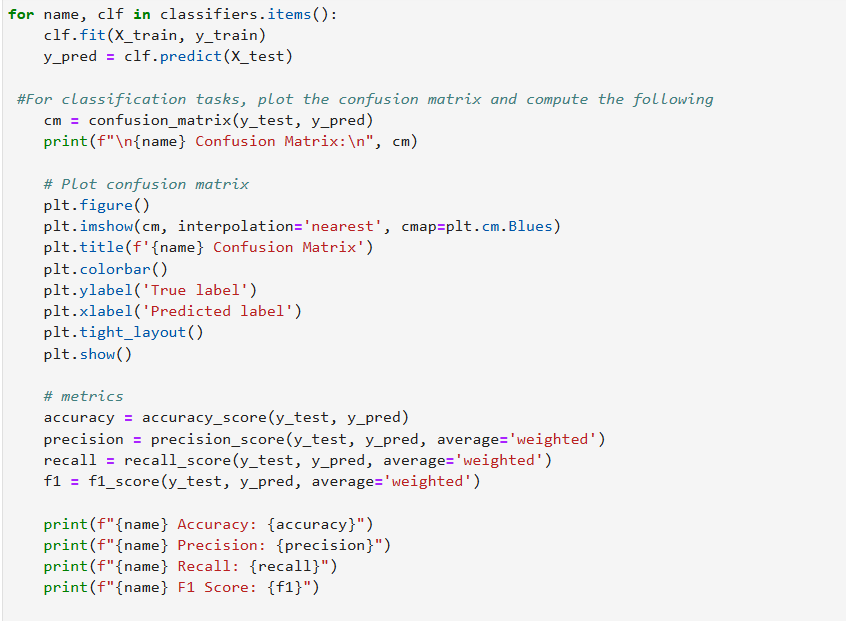
All the given task run successfully

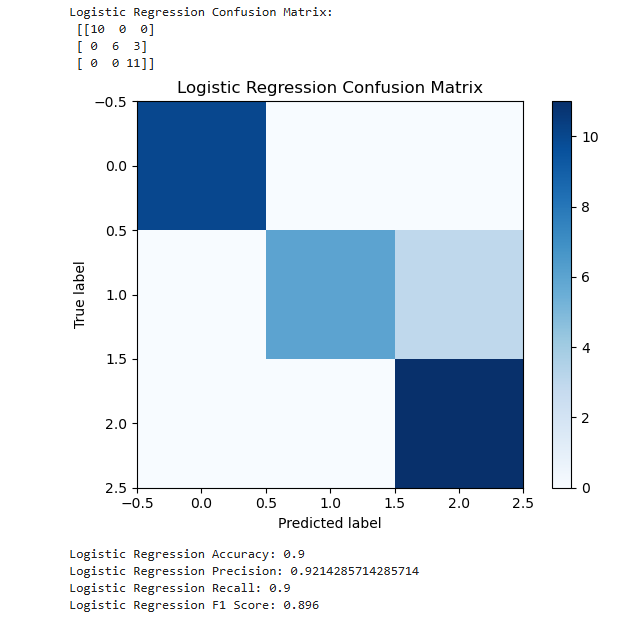
**Task5:**

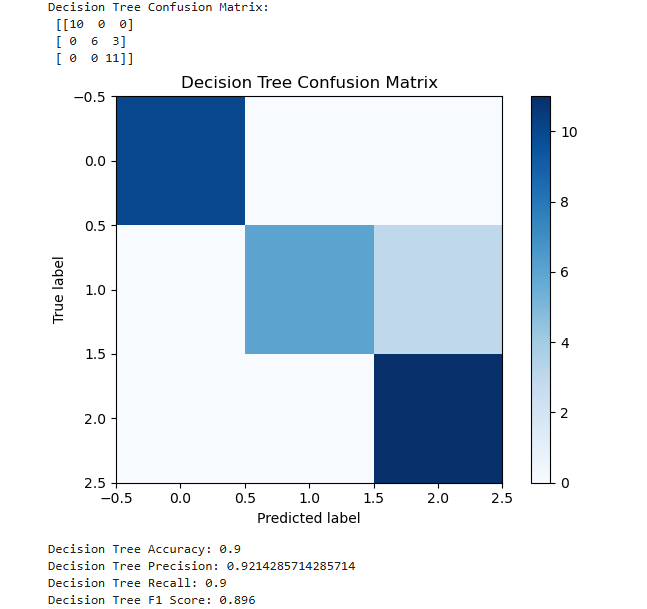
Confusion Matrix:

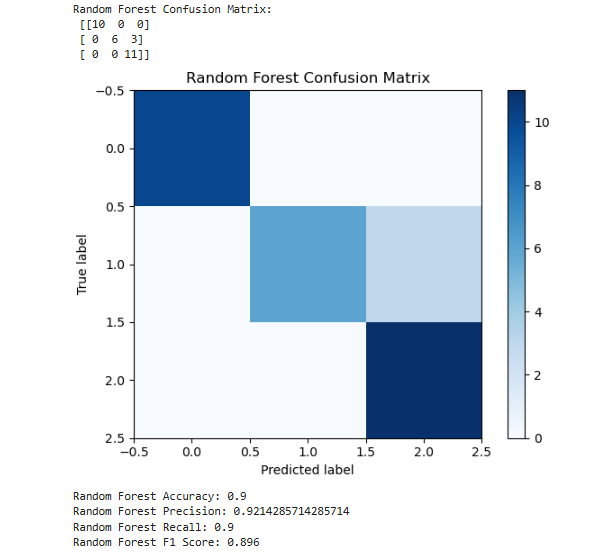
* For classification tasks, plot the confusion matrix and compute the following metrics:
* Accuracy
* Precision
* Recall
* F1 Score

**Code:**

****

****

****

****

**Result:**

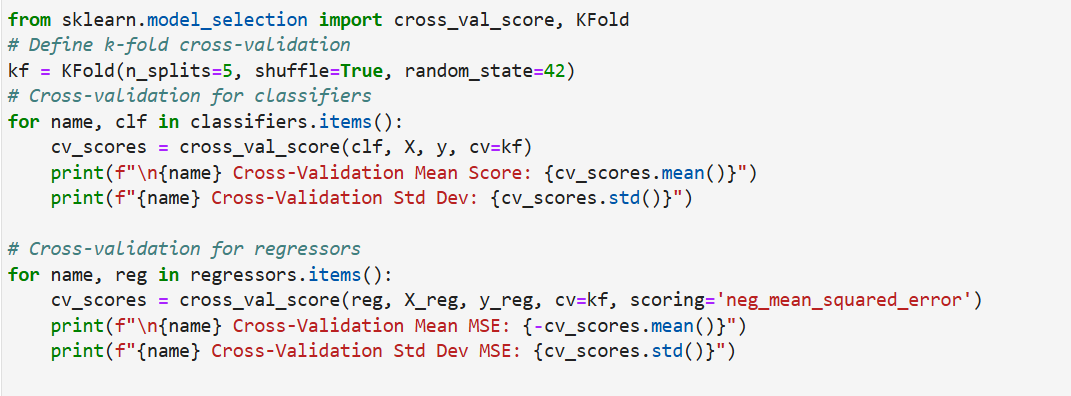
All the given task run successfully

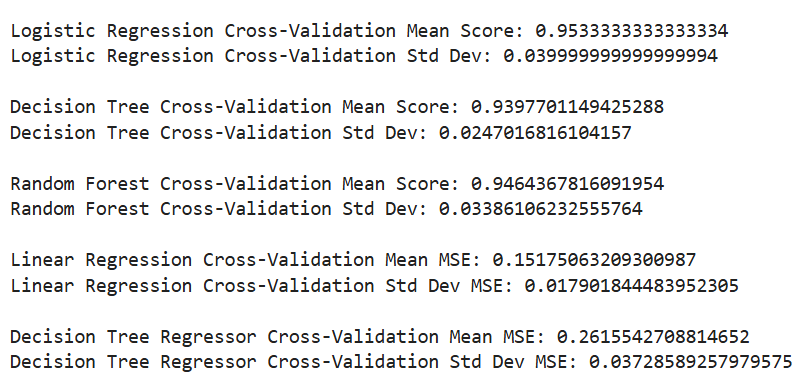
**Task6:**

Cross-Validation:

* Implement k-fold cross-validation for both classification and regression models.
* Report the mean and standard deviation of the cross-validation scores.

**Code:**





**Result:**

All the given task run successfully