

## Advance Track Capstone Project

**Problem Statement:** Build machine learning models for prediction. [Regression Task]

**Dataset:** House\_Price\_prediction.csv

### **Task 1: [20 Marks]**

- Analyse the dataset and perform the steps below to build **linear regression** machine learning model.
  1. Data Splitting – Split the data into training and testing datasets.
  2. Build the linear regression model using training dataset
  3. Make predictions using the built model on test dataset.
  4. Evaluate model using appropriate evaluation metrics. [MAE, RMSE, R-SQUARED SCORE]
- Add Python comments to explain your code blocks.

### **Task 2: [10 Marks]**

Perform the following operations:

1. Build the **linear regression model** by applying hyperparameter tuning using GridSearchCV.
2. Display the result of GridSearchCV for **linear regression** model such as model\_name, best\_score\_ and best\_params\_

**Problem Statement:** Build machine learning models for prediction. [Classification Task]

**Dataset:** cancer-data-2.csv

### **Task 3: [20 Marks]**

- Analyse the dataset and perform the steps below to build **k-Nearest Neighbour (kNN)** model.
  1. Data Splitting – Split the data into training and testing datasets.
  2. Build the classification model using training dataset
  3. Make predictions using the built model on test dataset.
  4. Evaluate model using appropriate evaluation metrics. [print confusion matrix and classification report]
- Add Python comments to explain your code blocks.

### **Task 4: [10 Marks]**

- Perform the following operations:
  1. Build **k-Nearest Neighbour (kNN)** classification model by applying hyperparameter tuning using GridSearchCV.

2. Display the result of GridSearchCV for **k-Nearest Neighbour (kNN)** model such as model\_name, best\_score\_ and best\_params\_

**Note –**

1. The comments and conclusions/observations carry marks for each part.
2. Students should submit their work in the form of .ipynb code file to Dr. Richa Dhanuka ([richad@regenesys.net](mailto:richad@regenesys.net))
3. Please follow the filename format below: Name\_of\_student-Batch\_name-Capstone\_Project\_2.ipynb.
4. Keep the same naming convention as mentioned above for email.
5. Last date of submission is **15 August 2024**.