Lab 2

Faris Alenezi – 2240004391

1- Dataset

Dataset Link: https://www.kaggle.com/datasets/juhibhojani/house-price

2- Defining the Machine Learning Problem

- We will use **regression** problem because when we deal with regression the goal is to predict a continuous numerical value which is (house price).
- **The target variable**: Price (in rupess).
- The model is expected to learn the **relationship between property features and house price features** such as:
    - Location
    - Carpet Area
    - Number of Bathrooms
    - Floor details
    - Furnishing status
    - Transaction type
    - Parking availability
    - Facing and overlooking

  And will predict the price of house in rupees.

- Problem Description:

  The goal of this machine learning is to build regression model that predicts house prices based on property like location, area, floor level and other housing features. By learning from housing data, the model aims to estimate the price of a house accurately.

3- Loading and Inspecting the Dataset in Python

- The Code:

lab2.py 1

C: > Users > Admin > Desktop > IAUL6 > Machine learning > LABS > 2 > lab2.py > ...

```python
import pandas as pd

df = pd.read_csv(r"C:\Users\Admin\Desktop\IAUL6\Machine learning\LABS\2\house_prices.csv")

print("Dataset Shape:")
print(df.shape)

print("\nFirst 5 Rows of the Dataset:")
print(df.head())

print("\nColumn Names:")
print(df.columns)

print("\nData Types of Each Column:")
print(df.dtypes)
```

- The Output:

```
PS C:\Users\Admin\Desktop\IAUL6\Machine learning\LABS\2> python .\lab2.py
Dataset Shape:
(187531, 21)

First 5 Rows of the Dataset:
   Index                                      Title                                        Description Amount(in rupees)  ...             Ownership Super Area Dimensions Plot Area
0      0  1 BHK Ready to Occupy Flat for sale in Srushti...  Bhiwandi, Thane has an attractive 1 BHK Flat f...          42 Lac   ...                   NaN        NaN        NaN       NaN
1      1  2 BHK Ready to Occupy Flat for sale in Dosti V...  One can find this stunning 2 BHK flat for sale...          98 Lac   ...              Freehold        NaN        NaN       NaN
2      2  2 BHK Ready to Occupy Flat for sale in Sunrise...  Up for immediate sale is a 2 BHK apartment in ...         1.40 Cr   ...              Freehold        NaN        NaN       NaN
3      3        1 BHK Ready to Occupy Flat for sale Kasheli  This beautiful 1 BHK Flat is available for sal...          25 Lac   ...                   NaN        NaN        NaN       NaN
4      4  2 BHK Ready to Occupy Flat for sale in TenX Ha...  This lovely 2 BHK Flat in Pokhran Road, Thane ...         1.60 Cr   ...  Co-operative Society        NaN        NaN       NaN

[5 rows x 21 columns]

Column Names:
Index(['Index', 'Title', 'Description', 'Amount(in rupees)',
       'Price (in rupees)', 'location', 'Carpet Area', 'Status', 'Floor',
       'Transaction', 'Furnishing', 'facing', 'overlooking', 'Society',
       'Bathroom', 'Balcony', 'Car Parking', 'Ownership', 'Super Area',
       'Dimensions', 'Plot Area'],
      dtype='str')

Data Types of Each Column:
Index                int64
Title                  str
Description            str
Amount(in rupees)      str
Price (in rupees)    float64
location               str
Carpet Area            str
Status                 str
Floor                  str
Transaction            str
Furnishing             str
facing                 str
overlooking            str
Society                str
Bathroom               str
Balcony                str
Car Parking            str
Ownership              str
Super Area             str
Dimensions           float64
Plot Area            float64
dtype: object
PS C:\Users\Admin\Desktop\IAUL6\Machine learning\LABS\2>
```

4- Designing the Methodology Diagram