HOUSE RENT PREDICTION

For house rent prediction using machine learning, the objective is to predict the rent of a house based on features such as the number of rooms (BHK), size, floor, area locality, furnishing status, and more. The steps to complete the machine learning project for house rent prediction are outlined below.

Steps for House Rent Prediction:

1. Problem Definition:

- **Objective:** Predict the rental price of a house based on features such as the number of rooms (BHK), area locality, size of the house, floor, furnishing status, etc.

- **Problem Type: Regression** (since the rent is a continuous value).

2. Collect the Dataset:

You can use datasets like Kaggle's House Rent Prediction Dataset or any other dataset that contains information on house rentals. A typical dataset may include columns like:

- BHK (Number of Bedrooms, Hall, Kitchen)

- Rent (Target variable)

- Size (Area in square feet)

- Floor (Floor number and total floors)

- Area Type (Super area or Carpet area)

- City

- Furnishing Status (Furnished, Semi-furnished, Unfurnished)

- Bathroom count

3. Preprocess the Data:

This step involves cleaning the data, handling missing values, encoding categorical variables, and scaling numerical features.

Load dataset:

Dataset = pd.read_csv('house_rent.csv')

Step 1: Handle missing values

Step 2: Label Encoding (for categorical columns like 'Area Type' or 'Furnishing Status')

Step 3: One-Hot Encoding (for categorical columns like 'City')

Step 4: Feature scaling (for continuous columns like 'Size', 'BHK', and ‘Rent’)

Step 5: Split the dataset into features (X) and target (y)

Step 6: Split data into training and testing sets

Step 7: Check preprocessed data

4.Exploratory Data Analysis[EDA]

-Uni-variate & bi-variate

5. Feature Selection

6. Model creation:

After preprocessing, you can proceed to select an appropriate regression model for house rent prediction. Some popular regression algorithms include:

- Linear Regression
- Decision Tree Regressor
- Random Forest Regressor
- XG-Boost Regressor

7. Model Evaluation:

- **Mean Absolute Error (MAE):** The average of the absolute errors.

- **Mean Squared Error (MSE):** The average of the squared errors.

- **R2 Score:** Represents how well the model explains the variance of the target variable. A score closer to 1 is ideal.

8. Hyper parameter Tuning:

- You can improve model performance by tuning the model's hyper parameters using **Grid Search** or Randomized Search.

9. Make Predictions:

- After training the model, you can use it to predict the rent of a house based on new data.