PREDICTING HOUSE PRICE USING MACHINE LEARNING

To start building a house price prediction model, you'll first need a dataset containing relevant information about houses and their prices. Common features include square footage, number of bedrooms and bathrooms, location, etc.

Assuming you have a dataset in a CSV format, I'll guide you through the steps of loading and preprocessing it using Python and popular libraries like pandas and scikit-learn.

**PROGRAM:**

# Import necessary libraries

import pandas as pd

from sklearn.model\_selection import train\_test\_split

from sklearn.preprocessing import StandardScaler

# Load the dataset

data = pd.read\_csv('house\_dataset.csv') # Replace 'house\_dataset.csv' with your actual file path

# Explore the dataset

print(data.head()) # Print the first few rows to understand the structure

# Separate features (X) and target (y)

X = data.drop('price', axis=1) # Assuming 'price' is the target variable

y = data['price']

# Split the dataset into training and testing sets

X\_train, X\_test, y\_train, y\_test = train\_test\_split(X, y, test\_size=0.2, random\_state=42)

# Standardize the features (optional but recommended for many models)

scaler = StandardScaler()

X\_train = scaler.fit\_transform(X\_train)

X\_test = scaler.transform(X\_test)

**Explanation:**

**1. Import Libraries:**

We start by importing the necessary libraries, including pandas for data handling and scikit-learn for machine learning functionalities.

**2. Load the Dataset:**

Use `pd.read\_csv()` to load your dataset. Replace `'house\_dataset.csv'` with the actual path to your dataset.

**3. Explore the Dataset:**

Print out the first few rows using `data.head()` to understand the structure and contents of your dataset.

**4. Separate Features and Target:**

In this example, it's assumed that 'price' is the target variable. We create `X` containing the features and `y` containing the target variable.

**5. Split the Dataset**:

We split the data into training and testing sets using `train\_test\_split`. This is crucial for evaluating the model's performance.

**6. Standardize the Features :**

Standardization can help improve the performance of some models. It ensures that all features are on a similar scale.

Now you have preprocessed data ready for building your prediction model. Depending on your preferences and the nature of the data, you can choose from various regression algorithms like Linear Regression, Random Forest, or even more complex models like Neural Networks for your house price prediction task. Let me know if you'd like guidance on building a specific type of model! # Import necessary libraries