**GOVERNMENT COLLEGE OF ENGINEERING ERODE**

****

B.E Electronics and Communication Engineering

PREDICTING HOUSE PRICES USING MACHINE LEARNING

**Name of the Students: University Register no:**

**Team Leader:**

M.Priyadharshini 731121106038

**Team Members:**

M.Sugi 731121106048

K.V.Hema 731121106017

Under the mentor of

**Dr.M.Sathyakala**

**Department of Information Technology(IT)**

**Department of Electronics and Communication Engineering**

Government College of Engineering

Near Vasavi College(PO),Erode,TamilNadu-638316,

Affiliated to Anna University ,Chennai.

**INTRODUCTION:**

* Predicting house prices using machine learning involves developing a model or algorithm that uses historical data and various features of a property to estimate its market value or selling price.

* House price prediction using machine learning involves using historical data and various algorithms to estimate the price of a house based on different features such as location, size, number of rooms, amenities, and more. Machine learning models like regression, decision trees, or neural networks can be trained on a dataset of past house prices and their corresponding features to learn patterns and make predictions on new, unseen data.
* The idea is to find the best model that can accurately predict house prices based on the given features. This can be useful for real estate agents, buyers, or sellers to get an estimate of a house's value in the market.

**PROJECT OVERVIEW:**

The overall, predicting house prices using machine learning involves a comprehensive data science pipeline, from data collection and preprocessing to model development, evaluation, deployment and ongoing maintenance.

**OBJECTIVE:**

* To develop a machine learning model that accurately predicts house prices based on various features.
* It is a helping tool for real estate professionals and individuals looking to buy or sell properties.

**DATASET:**

<https://www.kaggle.com/datasets/vedavyasv/usa-housing>

**Necessary step to follow:**

**1.Import Libraries:**

Start by importing the necessary libraries:

**Program:**

import pandas as pd

import numpy as np

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

from sklearn.preprocessing import StandardScaler

**2.Load the Dataset:**

Load your dataset into a Pandas DataFrame. You can typically find house price datasets in CSV format, but you can adapt this code to other formats as needed.

**Program:**

df = pd.read\_csv(' E:\USA\_Housing.csv ')

Pd.read()

**3. Exploratory Data Analysis (EDA):**

Perform EDA to understand your data better. This includes

checking for missing values, exploring the data's statistics, and

visualizing it to identify patterns.

**Program:**

# Check for missing values

print(df.isnull().sum())

# Explore statistics

print(df.describe())

# Visualize the data (e.g., histograms, scatter plots, etc.)

**4. Feature Engineering:**

Depending on your dataset, you may need to create new features or

transform existing ones. This can involve one-hot encoding categorical variables, handling date/time data, or scaling numerical features.

**Program:**

# Example: One-hot encoding for categorical variables

df = pd.get\_dummies(df, columns=[' Avg. Area Income ', ' Avg. Area

House Age '])

**5. Split the Data:**

Split your dataset into training and testing sets. This helps you evaluate your model's performance later.

**Program:**

X = df.drop('price', axis=1) # Features

y = df['price'] # Target variable

X\_train, X\_test, y\_train, y\_test = train\_test\_split(X, y, test\_size=0.2,

random\_state=42)

**6. Feature Scaling:**

Apply feature scaling to normalize your data, ensuring that all

features have similar scales. Standardization (scaling to mean=0 and

std=1) is a common choice.

**Program:**

scaler = StandardScaler()

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

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

**Importance of loading and processing dataset:**

* Loading and preprocessing the dataset is an important first step in building any machine learning model. However, it is especially important for house price prediction models, as house price datasets are often complex and noisy.
* By loading and preprocessing the dataset, we can ensure that the machine learning algorithm is able to learn from the data effectively and accurately.

**CONCLUSION:**

* In the quest to build a house price prediction model, we have

embarked on a critical journey that begins with loading and

preprocessing the dataset.We have traversed through essential

steps, starting with importing the necessary libraries to facilitate

data manipulation and analysis.

* Understanding the data's structure, characteristics, and any

potential issues through exploratory data analysis (EDA) is

essential for informed decision-making.

* Data preprocessing emerged as a pivotal aspect of this process.
* With these foundational steps completed, our dataset is now

primed for the subsequent stages of building and training a house

price prediction model.