**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>

**Preprocessing and Feature Scaling using Pipeline**

print("\n--- Feature Scaling ---")

model = Pipeline([ ('preprocessor', preprocessor), ])

# Fit the preprocessing pipeline on the training data

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

# Transform the testing data using the fitted pipeline

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

print("--- Preprocessing Complete! ---")

**Linear Regression:**

lin\_reg = LinearRegression()

lin\_reg.fit(X\_train, y\_train)

predictions = lin\_reg.predict(X\_test)

mae, mse, rmse, r\_squared = evaluation(y\_test, predictions)

print("MAE:", mae)

print("MSE:", mse)

print("RMSE:", rmse)

print("R2 Score:", r\_squared)

print("-"\*30)rmse\_cross\_val = rmse\_cv(lin\_reg)

print("RMSE Cross-Validation:", rmse\_cross\_val)

new\_row = {"Model": "LinearRegression","MAE": mae, "MSE": mse, "RM

SE": rmse, "R2 Score": r\_squared, "RMSE (Cross-Validation)": rmse\_cross \_val}models = models.append(new\_row, ignore\_index=True)

Split the data into training and test sets. The training set will be

used to train the model, and the test set will be used to evaluate

the performance of the model.

X\_train, X\_test, Y\_train, Y\_test = train\_test\_split(X, Y, test\_size=0.2, random\_state=101)

Y\_train.head()

**Model evaluation:**

There are a number of different metrics that can be used to evaluate

the performance of a house price prediction model. Some of the most

common metrics include:

• Mean squared error (MSE): This metric measures the average

squared difference between the predicted and actual house prices.

• Root mean squared error (RMSE): This metric is the square root

of the MSE.

• Mean absolute error (MAE): This metric measures the average

absolute difference between the predicted and actual house prices.

• R-squared: This metric measures how well the model explains the

variation in the actual house prices

**Evaluation of predicted data:**

plt.figure(figsize=(12,6))

plt.plot(np.arange(len(Y\_test)), Y\_test, label='Actual Trend')

plt.plot(np.arange(len(Y\_test)), Prediction5, label='Predicted Trend')

plt.xlabel('Data')

plt.ylabel('Trend')

plt.legend()

plt.title('Actual vs Predicted')

Model Comparison:

The less the Root Mean Squared Error (RMSE) is,

models.sort\_values(by="RMSE (Cross-Validation)")

**Conclusion:**

In summary, the successful prediction of house prices using machine learning involves a combination of data preparation, model selection, evaluation and on going maintenance.