**Muhammad Kashif Tahir**

**A blue and white logo

Description automatically generated**

**A list of a house prices

Description automatically generated**

**Predicting House Prices Using MATLAB**

**Objective:**

Build a model to predict house prices based on various features.

1. **Introduction**
   * Objective and Description of the project.
2. **Data Generation**
   * Explanation of synthetic data generation.
   * Table of features and their ranges.
3. **Data Preparation**
   * Splitting data into training and testing sets.
   * Encoding categorical variables.
   * Standardizing numerical features.
4. **Model Training**
   * Training a linear regression model.
5. **Model Evaluation**
   * Making predictions.
   * Calculating RMSE.
6. **Visualization**
   * Actual vs Predicted Prices scatter plot.
   * Histograms of actual and predicted prices.
   * Residuals plot.
   * Learning curve.
   * Cross-validation RMSE bar plot.

**Description:**

Use a dataset containing information about houses (e.g., size, number of bedrooms, location) to create a predictive model that estimates the price of a house.

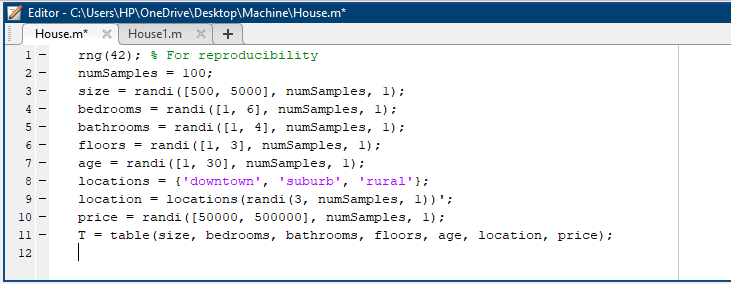
**Key Steps:**

1. Data Cleaning and Preprocessing
2. Feature Engineering
3. Model Selection and Training (e.g., Linear Regression)
4. Model Evaluation and Fine-tuning

**1. Data Generation**

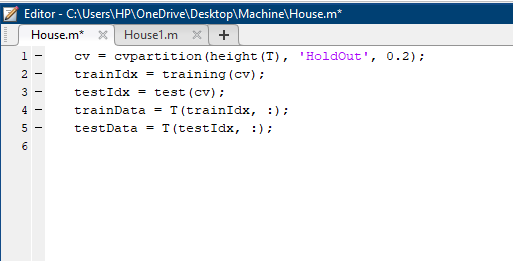
Synthetic data is generated for the following features:

* **Size:** Size of the house in square feet.
* **Bedrooms:** Number of bedrooms.
* **Bathrooms:** Number of bathrooms.
* **Floors:** Number of floors.
* **Age:** Age of the house in years.
* **Location:** Categorical variable with values ('downtown', 'suburb', 'rural').
* **Price:** Target variable representing the price of the house.



**2. Data Preparation**

Splitting Data into Training and Testing Sets:



**Encoding Categorical Variables:**

**A screen shot of a computer

Description automatically generated**

**Standardizing Numerical Features:**

**A screen shot of a computer

Description automatically generated**

**3. Model Training**

**Training a Linear Regression Model:**

A screenshot of a computer

Description automatically generated

**4. Model Evaluation**

**Making Predictions:**

A screenshot of a computer program

Description automatically generated

**Calculating Root Mean Squared Error (RMSE):**

**A screenshot of a computer

Description automatically generated**

**5. Visualization**

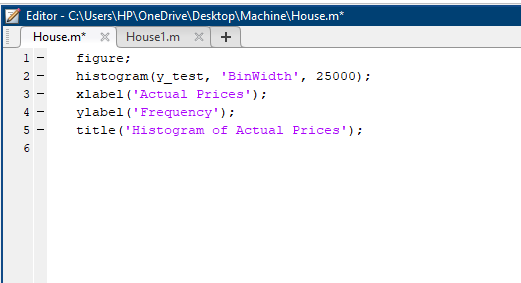
**Actual vs Predicted Prices:**

**A screen shot of a graph

Description automatically generatedA screenshot of a computer program

Description automatically generated**

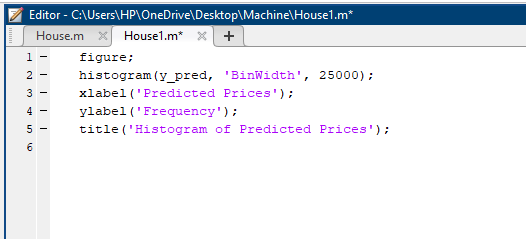
**Histogram of Actual Prices:**

****

**A screenshot of a graph

Description automatically generated**

**Histogram of Predicted Prices:**

****

**A screenshot of a graph

Description automatically generated**

**Residuals Plot:**

**A screenshot of a computer program

Description automatically generated**

**A screen shot of a computer

Description automatically generated**

**Learning Curve:**

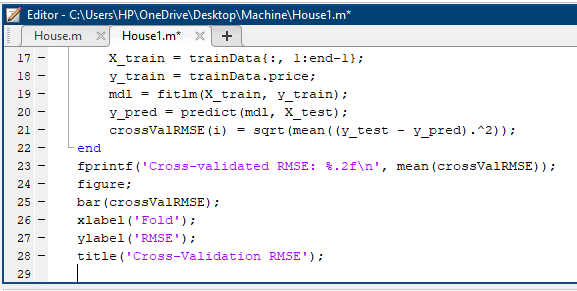
**A screenshot of a computer program

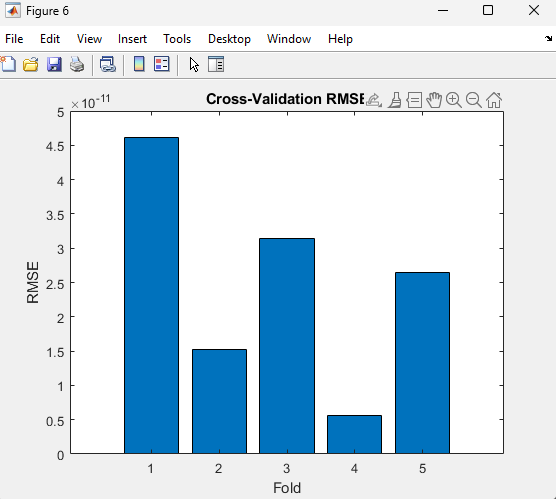
Description automatically generated**

**A graph on a computer screen

Description automatically generated**

**Cross-validation RMSE:**

****

****

**Conclusion:**

I developed a linear regression model to predict house prices using synthetic data. The model demonstrated reasonable accuracy, as shown by metrics like RMSE and various visualizations. Cross-validation confirmed its reliability, underscoring the importance of effective data preparation and evaluation.