**Linear Regression on California Housing Dataset**

This project demonstrates the use of **Linear Regression**, a fundamental machine learning algorithm, to predict housing prices based on various features of the **California Housing Dataset** provided by scikit-learn.

**🔍 Objective**

To build a model that predicts the **median house value** in California districts using features like:

* Median income
* Average number of rooms
* House age
* Population
* Latitude and Longitude, etc.

**🧮 Steps Involved**

1. **Data Loading**:  
   The dataset is fetched using fetch\_california\_housing() from sklearn.datasets.
2. **Data Preparation**:  
   Features (X) and target (y) are extracted and split into training and test sets using an 80/20 ratio.
3. **Model Training**:  
   A LinearRegression model is trained on the training data.
4. **Prediction & Evaluation**:  
   The model makes predictions on the test set. Evaluation is done using:
   * **Mean Squared Error (MSE)**
   * **R-squared Score (R²)**
5. **Visualization**:  
   A scatter plot of **actual vs predicted values** is generated to visually assess model accuracy. Points close to the diagonal line indicate accurate predictions.

**📊 Results**

* The model provides a **baseline regression performance**, capturing general trends in the data.
* While it may not be perfect due to the linear nature of the algorithm, it’s useful for interpretability and understanding feature impact.

**📌 Tools & Libraries**

* scikit-learn for ML
* pandas for data manipulation
* matplotlib for visualization