**House Price Prediction**

This project is focused on predicting real estate prices based on various features using machine learning models. The project uses Python libraries like Pandas, NumPy, Seaborn, Matplotlib, and scikit-learn to preprocess data, build models, and evaluate their performance.

**CODE**

[http://localhost:8888/notebooks/AI\_phase5.ipynb](#_top)

**DATA**

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

**TOOLS/IDE**

Anaconda Jupyter Notebook

**HOW TO USE IT**

To get started with this project, follow these steps:

1. Download the dataset ('USA\_Housing.csv') and place it in the same directory as the Python script.

2. Run the Python script to perform data analysis, model training, and evaluation.

**Project Structure**

The project is organized as follows:

• Importing Libraries: Import necessary Python libraries for data analysis and machine learning.

• Loading the Dataset: Read the 'USA\_Housing.csv' dataset into a Pandas DataFrame and display its information.

• Data Visualization: Visualize the dataset using various Seaborn plots and Matplotlib.

• Data Preprocessing: Prepare the data for model training by splitting it into features (X) and the target variable (Y). Perform feature scaling using StandardScaler.

• Linear Regression: Train and evaluate a Linear Regression model for price prediction.

• Support Vector Regression (SVR): Train and evaluate an SVR model for price prediction.

• Lasso Regression: Train and evaluate a Lasso Regression model for price prediction.

• Random Forest Regression: Train and evaluate a Random Forest Regression model for price prediction.

• XGBoost Regression: Train and evaluate an XGBoost Regression model for price prediction.

Results

Each model's performance is evaluated using three metrics: R-squared (r2\_score), Mean Absolute Error (MAE), and Mean Squared Error (MSE). The results are displayed and visualized for comparison.

**Conclusion**

This project demonstrates how to use various machine learning models to predict real estate prices based on different features. Each model's performance is evaluated, and the results are compared to determine which model performs the best for the given dataset.

Feel free to modify and experiment with different models and hyperparameters to further improve the predictions.