PREDICTING HOUSE PRICE USING MACHINE LEARNING

PROJECT :House price prediction



**INTRODUCTION:**

* In the initial phase of building the house price prediction model, the key steps involve loading and pre processing the data set.
* This process is fundamental for ensuring that the data is in a suitable format for analysis and modeling. By loading the housing data set, we aim to understand its structure and features
* Subsequently, pre processing steps such as handling missing values, scaling numerical features, and encoding categorical variables are essential to ensure data quality and uniformity.
* These preliminary actions pave the way for effective data analysis and the construction of an accurate prediction model.

CONTENT FOR PROJECT PHASE 2:

In this part you will begin building your project by loading and pre processing the data set. Start building the house price prediction model by loading and pre processing the data set. Load the housing data set and pre process the data.

DATA SOURCE:

Data set Link:**: [https://www.kaggle.com/datasets/vedavyasv/usa-housing](https://www.kaggle.com/datasets/vedavyasv/usa-housing" \t "https://survey.zohopublic.in/zs/_blank)**

**Data Collection and Pre processing:**  Importing the data set: Obtain a comprehensive data set containing relevant features such as square footage, number of bedrooms, location, amenities, etc. 

**Data pre processing:** Clean the data by handling missing values, outliers, and categorical variables. Standardize or normalize numerical features.

**Exploratory Data Analysis (EDA)**: Visualize and analyze the data set to gain insights into the relationships between variables. Identify correlations and patterns that can inform feature selection and engineering.Present various data visualizations to gain insights into the data set. Explore correlations between features and the target variable (house prices). Discuss any significant findings from the EDA phase that inform feature selection.

**Feature Engineering:** Create new features or transform existing ones to capture valuable information. Utilize domain knowledge to engineer features that may impact house prices, such as proximity to schools, transportation, or crime rates.Explain the process of creating new features or transforming existing ones.Showcase domain-specific feature engineering, such as proximity scores or composite indicators. Emphasize the impact of engineered features on model performance.

**Advanced Regression Techniques: **

**Ridge Regression**: Introduce L2 regularization to mitigate multi collinearity and over fitting. Lasso Regression: Employ L1 regularization to perform feature selection and simplify the model. 

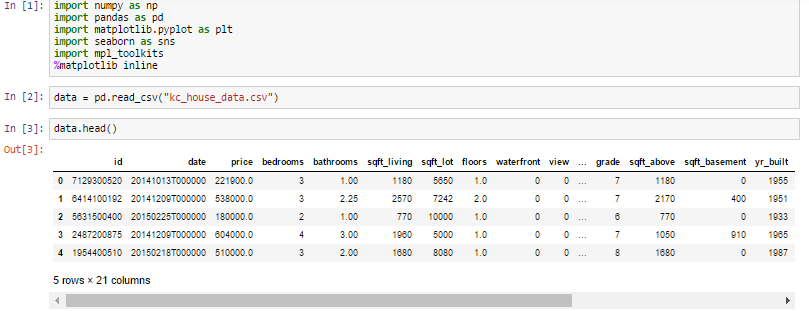
**Elastic Net Regression:** Combine both L1 and L2 regularization to benefit from their respective advantages. 

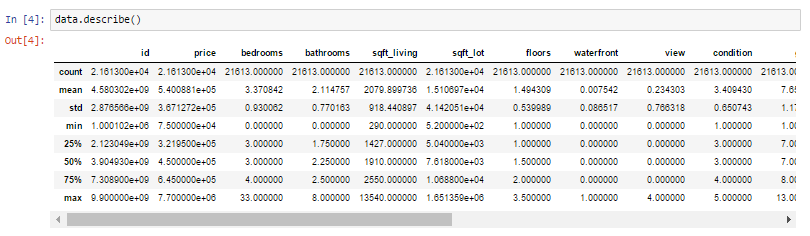
**Random Forest Regression:** Implement an ensemble technique to handle non-linearity and capture complex relationships in the data.  Gradient Boosting Regressors (e.g., XGBoost, LightGBM): Utilize gradient boosting algorithms for improved accuracy.

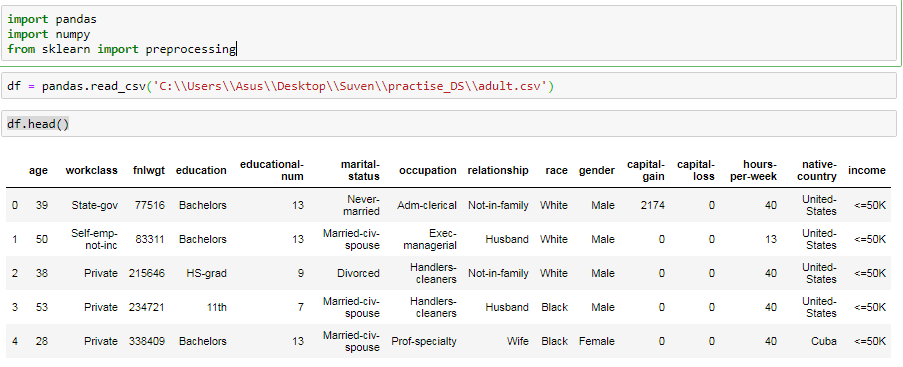
**Deployment and Prediction:**  Deploy the chosen regression model to predict house prices.  Develop a user-friendly interface for users to input property features and receive price predictions.

**PROGRAM:**

**HOUSE PRICING DETECTION**

****

****

****

**Conclusion and Future Work (Phase 3):**

**Project Conclusion:** 

In the Phase 3 conclusion, we will summarize the key findings and insights from the advanced preprocess techniques. We will reiterate the impact of these techniques on improving the accuracy and robustness of house price predictions. 

**Future Work:**

We will discuss potential avenues for future work, such as incorporating additional data sources (e.g., real-time economic indicators), exploring deep learning models for prediction, or expanding the project into a web application with more features and interactivity.