Project Title

Predicting Residential Property Prices Using Linear Regression in Python

Objective

Develop a predictive model to estimate house prices based on housing features such as area, bedrooms, furnishing status, and more.

Dataset Overview

Dataset: Housing.csv. Key columns include: Area, Bedrooms, Bathrooms, Furnishing Status, Parking, and Price (target variable).

Data Preprocessing

Data was loaded using pandas. Missing values and duplicates were handled. Categorical features were encoded using one-hot encoding. The dataset was split into training and test sets using train_test_split.

Exploratory Data Analysis

Matplotlib and Seaborn were used to visualize data. Plots included price distributions, area vs price scatter plots, and a correlation heatmap to understand feature relationships.

Model Used

Linear Regression was used to build a model assuming a linear relationship between features and the target house price.

Model Evaluation

Model was evaluated using R² Score (around 0.80), and RMSE for error analysis. A scatter plot of actual vs predicted values showed good alignment.

Conclusion

The model performed well with reasonable accuracy and interpretability. It's suitable as a baseline model for house price prediction tasks.

Future Enhancements

Future work could explore polynomial regression, random forests, or gradient boosting models. Adding more features such as location coordinates or year-built could further improve performance.

Tools & Technologies

Python, Pandas, NumPy, Scikit-learn, Matplotlib, Seaborn, Jupyter Notebook