

Linear Regression Report

Linear Regression Report

Title

Predicting House Prices using Linear Regression

1. Introduction

This project applies both simple and multiple linear regression techniques to a housing dataset to predict property prices based on various features. The implementation uses Python's Scikit-learn, Pandas, Matplotlib, and Seaborn libraries.

2. Dataset Description

Dataset: Housing.csv

Features include:

- Area (sqft)
- Number of Bedrooms
- Number of Bathrooms
- Stories
- Mainroad access
- Guestroom availability
- Basement availability
- Hot water heating
- Air conditioning

Linear Regression Report

- Parking
- Furnishing status
- Price (target variable)

Categorical variables were encoded using one-hot encoding.

3. Methodology

3.1 Preprocessing

- Categorical variables were encoded using Pandas `get_dummies()`.
- Data was split into training (80%) and test (20%) sets using `train_test_split`.

3.2 Model Training

- A `LinearRegression` model from Scikit-learn was trained on the training set.

3.3 Model Evaluation

Metrics used:

- Mean Absolute Error (MAE)
- Mean Squared Error (MSE)
- R-squared (R^2)

3.4 Visualization

- Actual vs Predicted price plot using Seaborn scatter plot.
- Regression line plotted for visual interpretation.

Linear Regression Report

4. Results

Evaluation Metrics:

Metric	Value
--------	-------

-----	-----
-------	-------

MAE	~value
-----	--------

MSE	~value
-----	--------

R ²	~value
----------------	--------

(Note: Insert actual output values here after execution)

Coefficient Interpretation

Each coefficient represents the effect of a feature on house price, holding other features constant. Positive values increase price prediction, negative values reduce it.

5. Insights & Discussion

- R² Score showed that a significant portion of variance in housing prices is explained by the model.
- Features like location (mainroad, furnishing status), air conditioning, and area had strong influence on prices.
- Residuals appeared randomly distributed, supporting the assumption of linearity.

6. Limitations & Future Work

- Assumes linear relationships; non-linear models may provide better performance.
- Could benefit from feature selection or regularization (e.g., Ridge, Lasso).

Linear Regression Report

- Might consider using cross-validation and hyperparameter tuning.

7. Conclusion

Linear regression is an interpretable and effective model for understanding the relationship between house features and price. This project reinforced core regression concepts and metrics, setting the stage for more advanced modeling.

End of Report