
AIRBNB PRICE PREDICTION.



RECAP

Overview: What triggers the Airbnb rental price?

Solution: Analyse the factors that influence the rental price and develop models that captures the complexity of the pricing system.

Impact: Provides interesting insights that can benefit a host looking to maximize their profit.

DATASET AND PREPROCESSING.

Dataset:

- Comprises information on Airbnb listings in USA with **74111 rows** and **29 columns**

Preprocessing:

- **Data cleaning:** Almost clean !
- **Scaling:** Scale numerical features to similar range.
- **Encoding:** Label encoding.
- **Feature selection** : Identify the key features that are likely to influence.

IMPORTANT FINDINGS IN EDA.

- **Data Distribution:** Normal distribution.
- **Data Quality:** Missing values and outliers were handled appropriately.
- **Patterns and trends:** To understand underlying relationships and phenomena of data through visualization.
- **Correlation:** Analyze multicollinearity.
- **Feature importance:** Determining which features are most relevant for the target variable through visual inspection.

BASELINE MODEL AND EVALUATION METRICS.

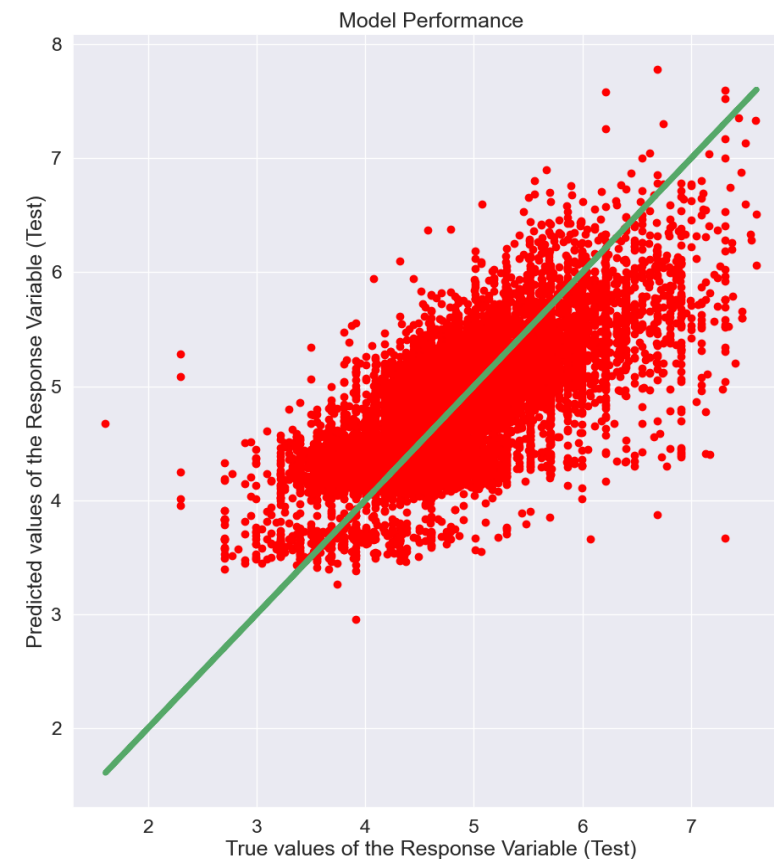
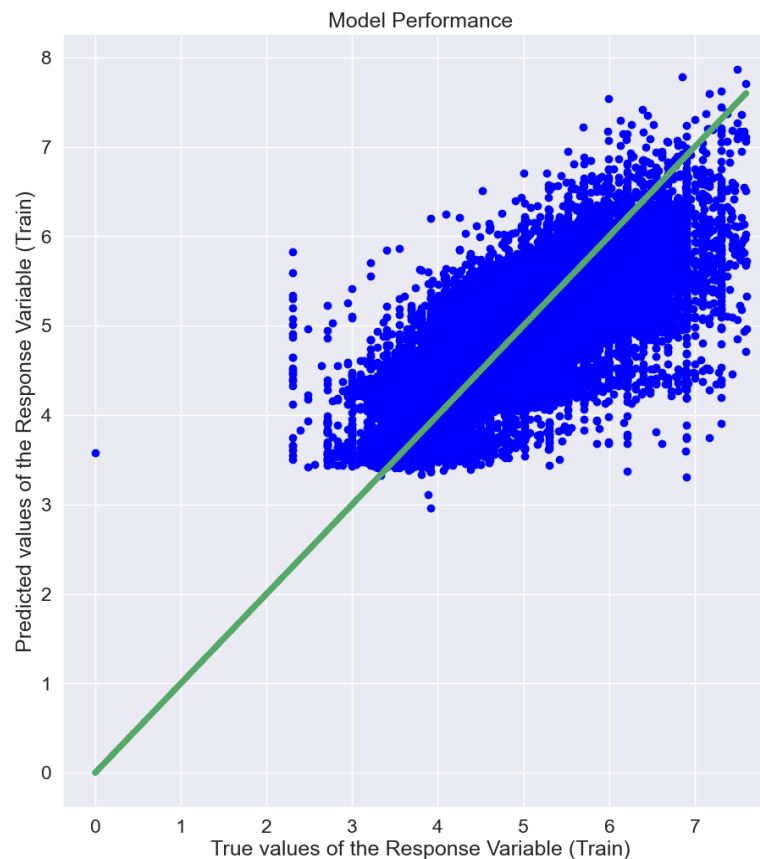
Baseline Model : **Linear
Regression**

Mean Absolute Error (MAE): **0.36**

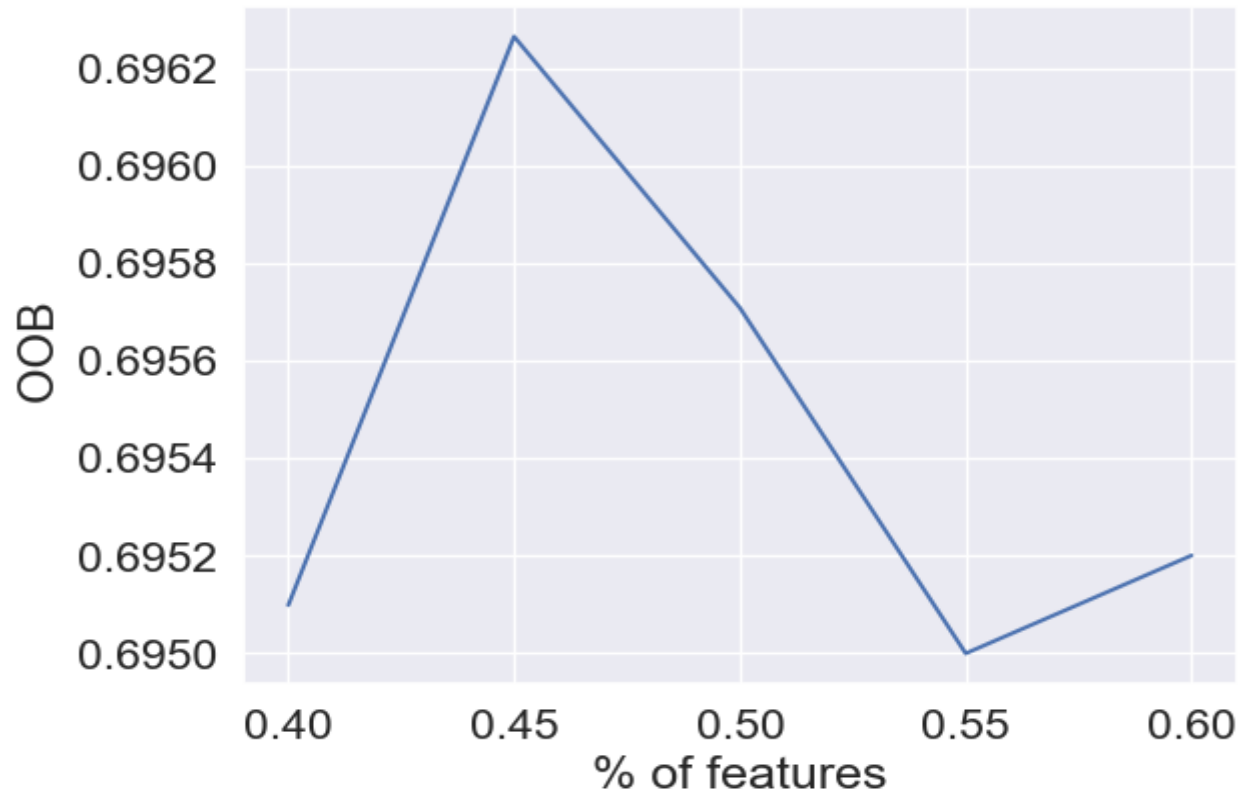
Mean Squared Error (MSE): **0.23**

Root Mean Squared Error
(RMSE): **0.48**

R-squared (R²) Score: **0.54**



RANDOM FOREST REGRESSOR



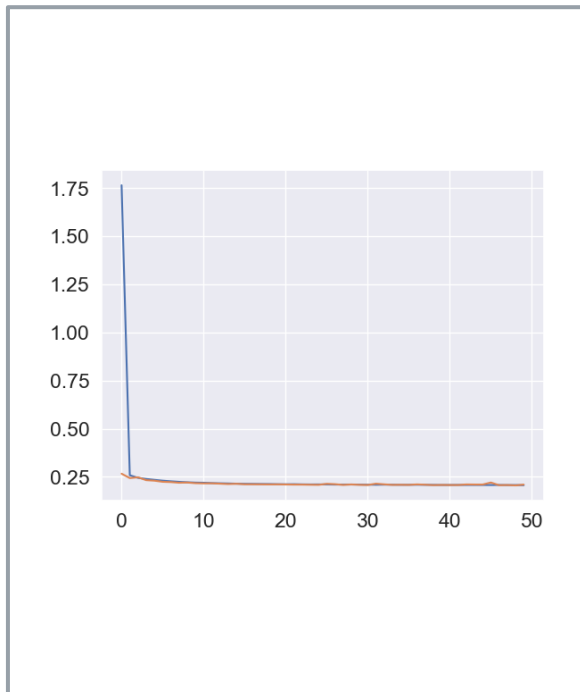
OOB Score: 70%

	feature	importance
1	room_type	0.279940
14	longitude	0.124422
13	latitude	0.106926
17	bedrooms	0.100541
3	accommodates	0.084841
4	bathrooms	0.063458

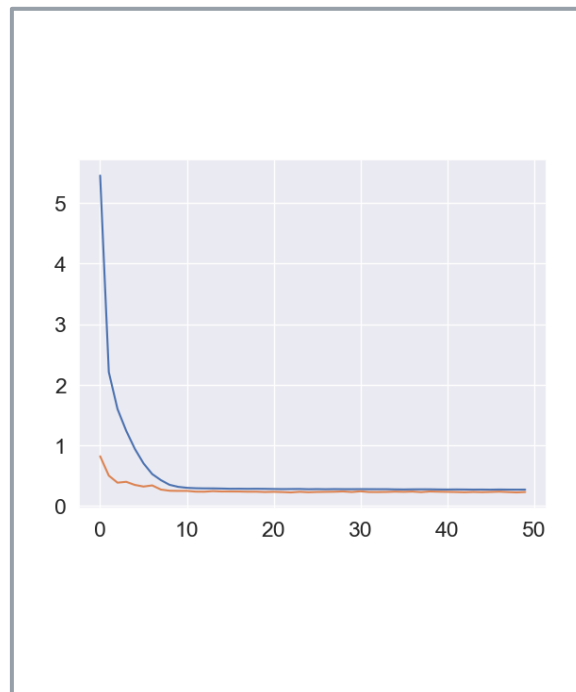
Importance

NEURAL NETWORK

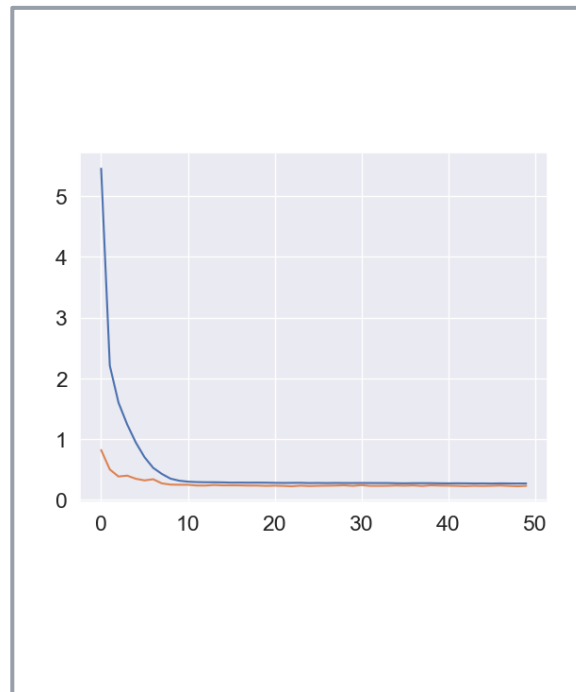
No feature selection
MAE: 0.33



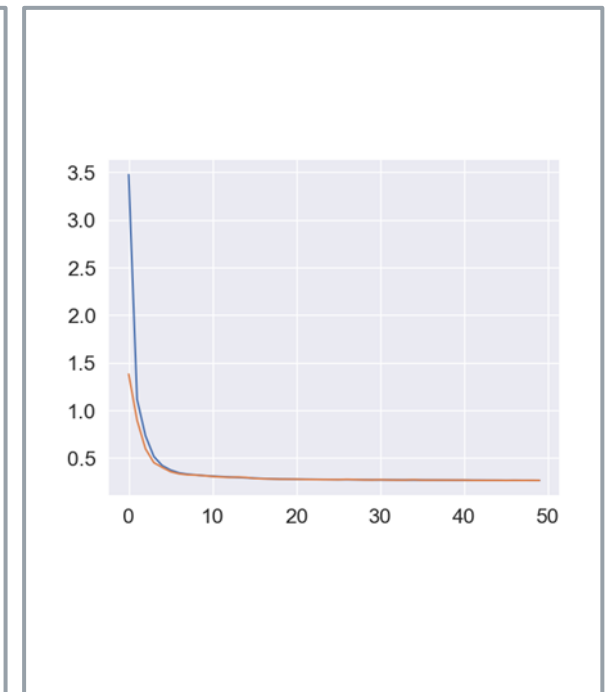
With PCA
MAE: 0.34



With Drop-out
MAE: 0.36



With Regularisation
MAE: 0.33





COMPARISON OF MODELS

Linear Regression: 54%

Random Forest: 70%

L1 Regularization: MAE: 0.33

THANK YOU!

ANY QUESTIONS?

