

HOUSING PRICES PREDICTION

Sang Uk Park

Brown University - Data I050

<https://github.com/sangupark?tab=repositories>

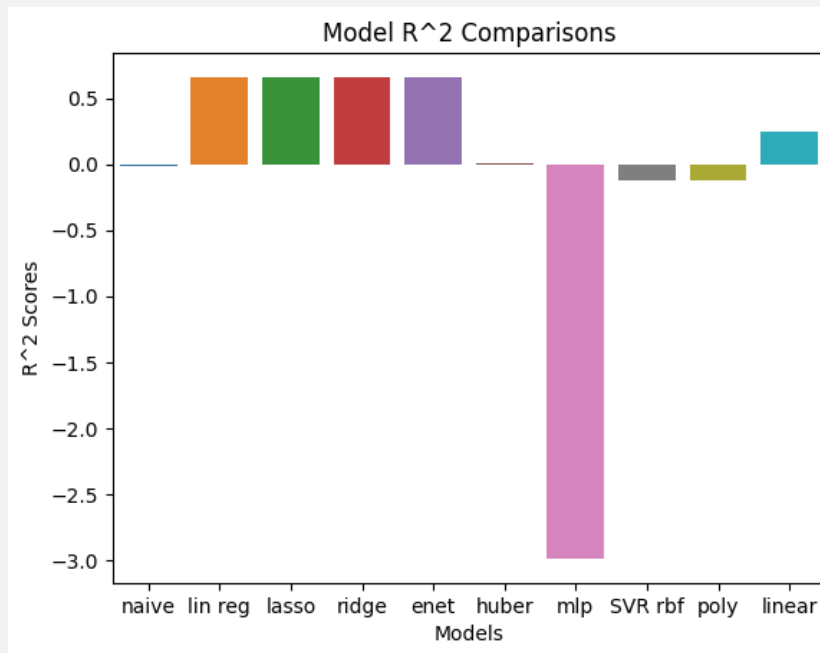
RECAP

- Housing Prices Data Set by M Yasser H
 - From Kaggle
- 12 Features
- Regression Model
- Interesting Finds
 - Price Distribution

CROSS-VALIDATION

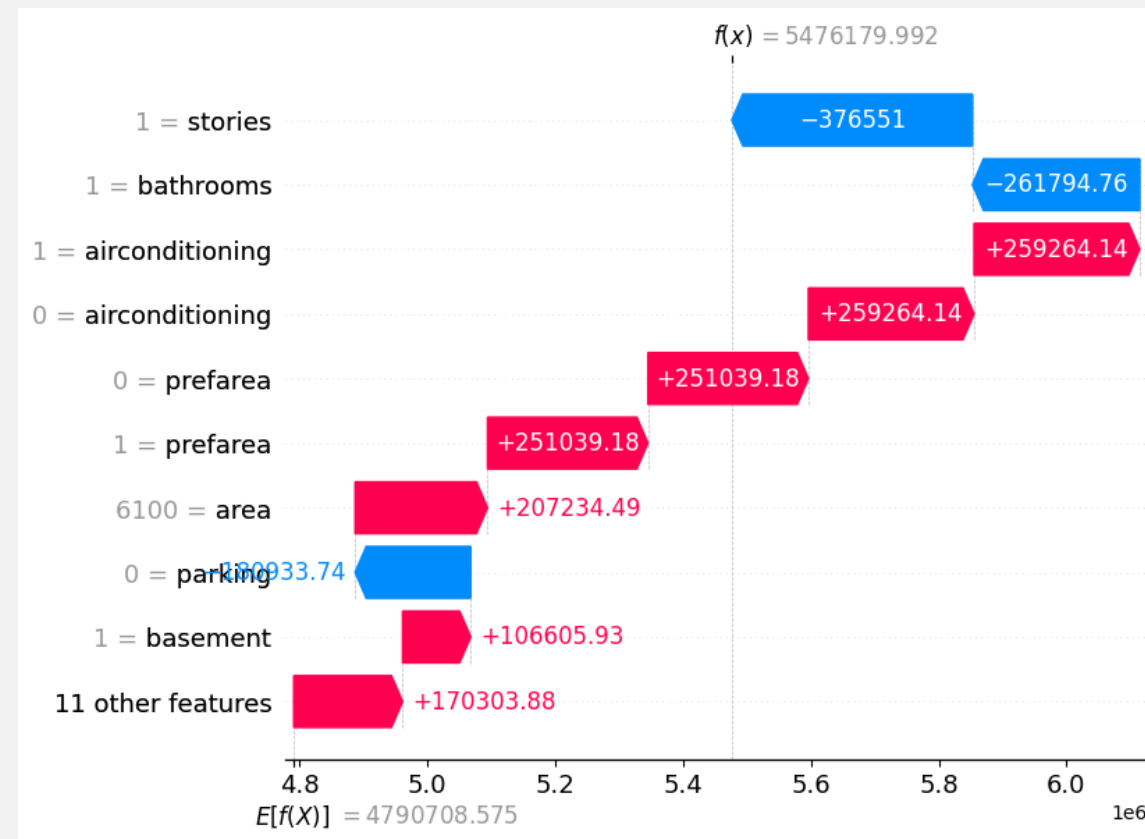
- Shuffled, 80% Training, 20% Testing
- ML Algorithms Tried
 - Linear Regression
 - Lasso
 - Ridge
 - Elastic Net
 - Huber Regressor
 - Multi-Layer Perceptron Regressor

RESULTS



area	0.15996	+/-	0.03125
bathrooms	0.15258	+/-	0.02243
stories	0.07344	+/-	0.02531
airconditioning	0.03008	+/-	0.00878
basement	0.00775	+/-	0.00235
guestroom	0.00698	+/-	0.00238

RESULTS CONTD.



RESULTS CONTD.

- Most Important Features
 - Area
 - Bathroom
 - Stories
 - AC
- Linear Models outperform Non-linear Models

OUTLOOK

- Dataset might not be representative enough
 - Might need more Segmentation
- Would be interested to see how different cities affect the pricing