House Price Analysis Report

Abstract

This report analyzes the impact of various features on house prices using Ridge and Lasso regression models. The analysis identifies the top 10 features influencing house prices and provides insights into how specific characteristics can affect a house's value.

Objective

To determine the most significant factors affecting house prices and quantify their impact using Ridge and Lasso regression techniques.

Introduction

In the real estate market, understanding the factors that influence house prices is crucial for buyers, sellers, and investors. This study employs two popular regularization techniques in machine learning - Ridge and Lasso regression - to identify and rank the most important features affecting house prices.

Methodology

1. Data Import
2. Data Cleaning
3. Data preprocessing and feature selection
4. Application of Ridge regression model
5. Application of Lasso regression model
6. Comparison of top 10 features from both models
7. Interpretation of coefficients to infer price impacts

Code

While the full code is not visible in the image, I'm attaching drive link of python notebook here:

https://drive.google.com/drive/u/1/folders/1_b4ldzCloIWUjIGRGCDaIUpTCsCDaI-V

Conclusion

Top Features (Ridge Model):

- OverallQual_9   1.133088
- Neighborhood_StoneBr   1.098437
- OverallQual_8   1.088531
- Neighborhood_Crawfor   1.087379
- Exterior1st_BrkFace   1.086563
- Neighborhood_NridgHt   1.084757
- LandContour_HLS   1.074319
- CentralAir_Y   1.072621
- OverallCond_9   1.072471
- BsmtCond_TA   1.068905

Top Features (Lasso Model):

- OverallQual_9   1.195210
- GrLivArea   1.116481
- Neighborhood_Crawfor   1.103855
- OverallQual_8   1.103809
- Neighborhood_StoneBr   1.093235
- Neighborhood_NridgHt   1.091923
- Exterior1st_BrkFace   1.087252
- CentralAir_Y   1.081540
- BsmtCond_TA   1.074009
- Functional_Typ   1.073811

Key Inferences:

1. The price of a house will increase by 1.11 times with an increase in GrLivArea (likely referring to Greater Living Area).
2. If the overall finish of the house is Very Good, the price can increase by 1.08 times..
3. Houses with centralized AC can see a price increase of 1.08 times.
4. A typical basement condition can increase the house price by up to 1.06 times.
5. The neighborhood has a significant impact on price, with Crawford, Stone Brook, and Northridge Heights being particularly desirable locations.

Both models agree on the importance of overall quality, specific neighborhoods, and exterior features in determining house prices. The Lasso model also highlights the significance of the living area size. These insights can be valuable for real estate professionals, homeowners, and potential buyers in understanding and predicting house values.