

Capstone Project 3 Proposal

House Prices: Advanced Regression Techniques

Predict sales prices and practice feature engineering, regression and ensemble methods.

Background

Ask a home buyer to describe their dream house, and they probably won't begin with the height of the basement ceiling or the proximity to an east-west railroad. But this playground competition's dataset proves that much more influences price negotiations than the number of bedrooms or a white-picket fence.

With 79 explanatory variables describing (almost) every aspect of residential homes in Ames, Iowa, this competition challenges you to predict the final price of each home.

<https://www.kaggle.com/c/house-prices-advanced-regression-techniques>

Goals

People always look for many features before buying their dreamed comfortable house, as it is a lifetime investment, by both Money and time. From the available dataset I would like to analyze to find out the following.

- What are the most features people are looking for
- Is there any correlation between the above features and the sale price
- Is it possible to predict sale price ?

Data Collection

The Ames Housing dataset (<http://www.amstat.org/publications/jse/v19n3/decock.pdf>) was compiled by Dean De Cock for use in data science education. It's an incredible alternative for data scientists looking for a modernized and expanded version of the often cited Boston Housing dataset.

Approach

1. Understand the variables in the data
2. Clean data to exclude redundant entries
3. EDA

4. Construct models with various techniques

5. Find out the output

Deliverables

1. Report/Paper

2. Jupyter Notebooks with data analysis

3. Model Accuracy Details