House Prices: Advanced Regression Techniques

Kaggle Competition Analysis

Zaeem Sabir

ROll No # 2024-MS-DS-136

Competition Overview

Objective

Predict house sale prices using advanced regression techniques.

Dataset

79 features describing various aspects of residential homes in Ames, Iowa.

Problem Statement

Predict the final sale price of homes based on provided features.

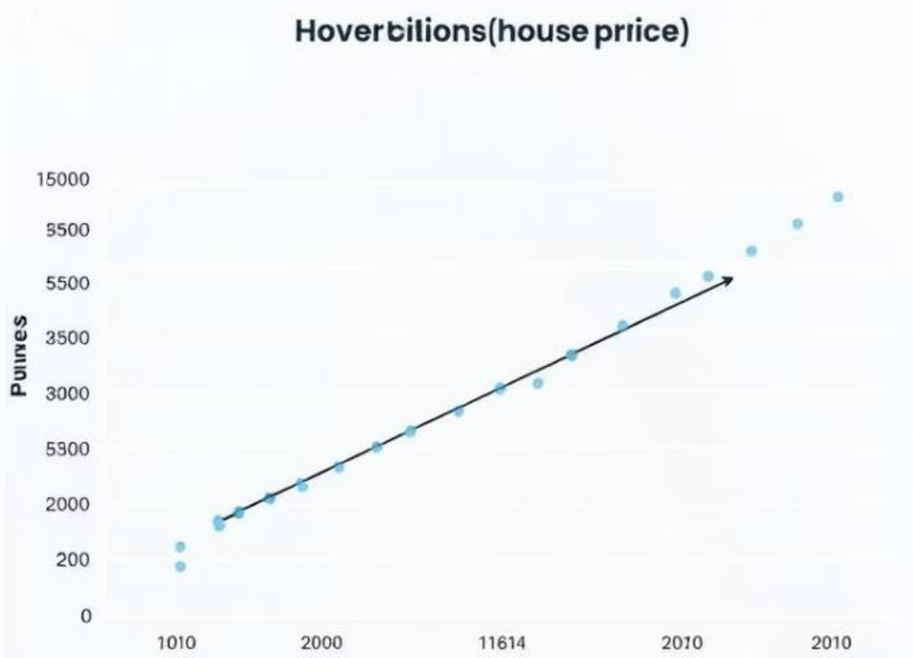
Data Exploration

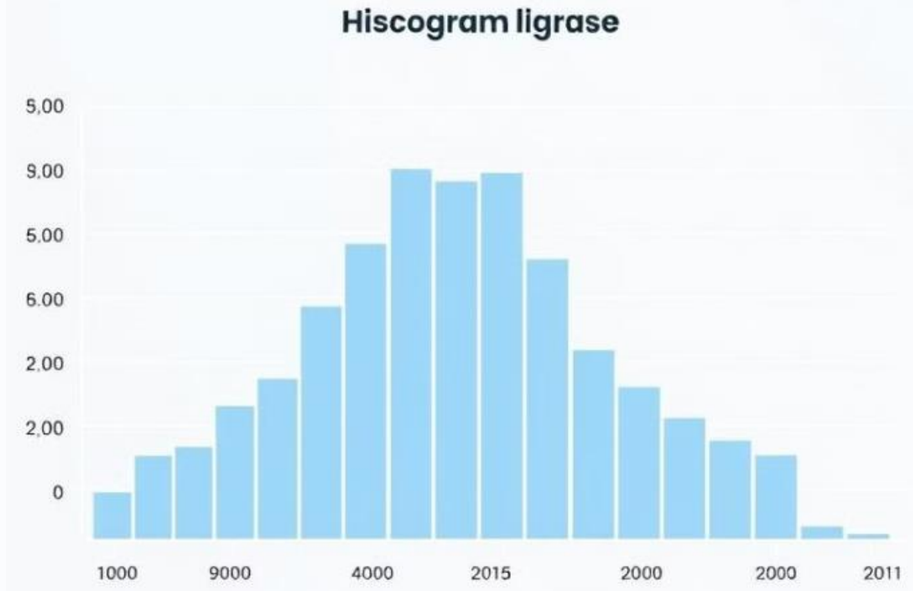
1 Initial Analysis

Understand data characteristics, identify potential trends, and spot outliers.

2 Missing Values

Address missing data points using imputation or feature removal.





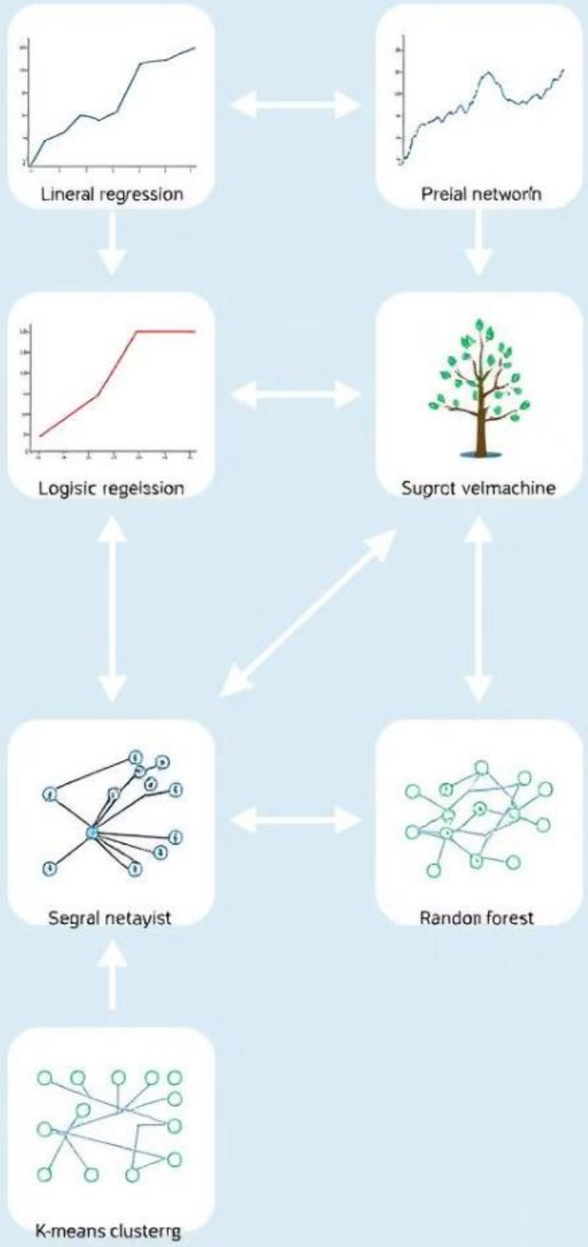
Feature Engineering

Feature Selection

Select features most relevant to predicting house prices.

Feature Transformation

Encode categorical variables, scale numerical features, and engineer new features.



Modeling

Model Selection

Explore linear regression, random forests, gradient boosting, and other regression models.

Model Training

Train models on the dataset, using techniques like train-test split and cross-validation.

Model Evaluation

Performance Metrics

Evaluate model performance using MSE, RMSE, and other relevant metrics.

Results

Compare the performance of different models and identify the best-performing one.

Conclusion

1

Findings

Summarize key insights from data exploration and model evaluation.

2

Future Work

Suggest potential improvements or next steps, such as exploring additional features or advanced techniques.