

# PREDICTING HOUSE PRICES USING MACHINE LEARNING

## INTRODUCTION

Predicting house prices is an important objective. Recent studies have found that asset prices can help to forecast output and inflation. Changes in housing prices can provide knowledge about consumption and inflation. In the business cycle, housing market play an important role. Studying Housing sales and prices is necessary to understand demand and supply. Models that can forecast housing prices can provide policy makers and economists with an idea about the future and therefore, help them design better policies. It can also help economists to estimate prepayments, housing mortgage and housing affordability. In recent years, machine learning has become a popular approach to predict house prices based on its attributes. Predicting house prices is usually approached as a regression problem. While traditional linear regression models can provide reasonable predictions, advanced regression techniques such as Gradient Boosting and XG Boost offer improved prediction accuracy and robustness. In this

article, we will explore how these advanced techniques can enhance the accuracy of house price predictions.

## **ADVANCED REGRESSION TECHNIQUES**

Advanced regression techniques used for house price prediction offer improved

accuracy and flexibility compared to traditional linear regression models. Some of

these

techniques are,

- Gradient Boosting Regression
- Random Forest Regressor
- Support Vector Regression (SVR)
- Lasso and Ridge Regression
- Elastic Net Regression
- Neural Networks (Deep Learning)

# GRADIENT BOOSTING

## REGRESSION

Gradient Boosting is an ensemble learning technique that builds a predictive model in the form of an ensemble of decision trees.

### TECHNIQUES:

- **XGBoost:** Extreme Gradient Boosting is a popular variation known for its efficiency and effectiveness in handling complex data.
- **LightGBM:** Light Gradient Boosting Machine is another efficient implementation that uses histogram-based learning.
- **CatBoost:** Categorical Boosting is designed to handle categorical features seamlessly

# **RANDOM FOREST**

## **REGRESSION**

Random Forest is an ensemble method that builds multiple decision trees and combines their predictions

### **TECHNIQUES:**

- Feature Importance: Random Forest can help identify important features for predicting house prices.
- Hyperparameter Tuning: Fine-tune parameters like the number of trees and tree depth for optimal performance.

# SUPPORT VECTOR REGRESSION (SVR))

SVR is a regression technique that uses support vector machines to find the hyperplane that best fits the data.

## TECHNIQUES:

- Kernel Functions: Choose appropriate kernel functions (e.g., linear, polynomial, radial basis function) based on data characteristics.
- Parameter Tuning: Adjust hyperparameters like the kernel coefficient and regularization term.

# LASSO AND RIDGE REGRESSION

Lasso and Ridge Regression are regularization techniques that help prevent overfitting in linear regression models.

## TECHNIQUES:

- L1 and L2 Regularization: These techniques introduce penalty terms to control the complexity of the regression model.
- Cross-Validation: Use cross-validation to choose the optimal regularization strength.

# ELASTIC NET

## REGRESSION

Elastic Net combines L1 (Lasso) and L2 (Ridge) regularization to balance feature selection and model simplicity.

### TECHNIQUES:

- Finding the Optimal Mix: Tune the mixing parameter between L1 and L2 regularization to achieve the desired balance.

# NEURAL NETWORKS

## (DEEP LEARNING)

Deep Learning models, particularly neural networks, can be applied to regression tasks, including house price prediction.

### TECHNIQUES:

- Architecture Design: Experiment with different neural network architectures, including feedforward, convolutional, or recurrent networks.
- Regularization: Apply dropout, batch normalization, and weight regularization to prevent overfitting.



# CONCLUSION

In summary, advanced regression techniques for house price prediction encompass a wide range of methods, from ensemble learning with Gradient Boosting to regularization methods like Lasso and Ridge Regression. The choice of technique depends on the characteristics of your dataset and the specific problem you are addressing. Experimentation and thorough evaluation are key to finding the most effective approach

# THANK YOU !!

