# Random Forest Regression and AdaBoost Classification

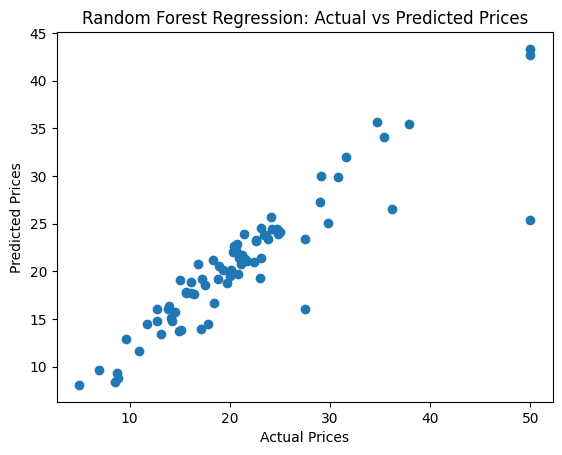
## Introduction

This document presents the implementation of Random Forest regression and AdaBoost classification in Boston Housing dataset and Breast Cancer dataset respectively.

## 

## Regression using Random Forests

The **Mean Squared Error (MSE):** 15.74



Model Training and Hyperparameter Tuning:

The Random Forest Regression model was trained using the Boston Housing dataset.

Hyperparameters were tuned using GridSearchCV, optimizing the number of trees and maximum depth.

The best model achieved an MSE of 15.74 on the test set.

Evaluation:

The model's performance was evaluated using the test set.

The scatter plot of predicted versus actual prices shows a moderately strong correlation.

## Classification using AdaBoost

* **Custom AdaBoost Classification (Breast Cancer Dataset)**

Classification report:

Custom AdaBoost Classification Report:

precision recall f1-score support

0 0.92 0.85 0.88 26

1 0.94 0.97 0.95 60

accuracy 0.93 86

macro avg 0.93 0.91 0.92 86

weighted avg 0.93 0.93 0.93 86

Confusion Matrix:

[[22 4]

[ 2 58]]

Model Training and Hyperparameter Tuning:

The Custom AdaBoost Classification model was trained on the Breast Cancer dataset.

Hyperparameters were tuned using GridSearchCV, optimizing the number of weak learners.

The best model achieved an accuracy of 93% on the test set.

Evaluation:

The model's performance was evaluated using the test set.

Precision, recall, and F1-score metrics indicate strong performance in classifying breast cancer samples.

**Sklearn AdaBoost Classification (Breast Cancer Dataset)**

Classification report:

Sklearn AdaBoost Classification Report:

precision recall f1-score support

0 0.96 1.00 0.98 26

1 1.00 0.98 0.99 60

accuracy 0.99 86

macro avg 0.98 0.99 0.99 86

weighted avg 0.99 0.99 0.99 86

Confusion Matrix:

[[26 0]

[ 1 59]]

Model Training and Hyperparameter Tuning:

The Sklearn AdaBoost Classification model was trained on the Breast Cancer dataset.

Hyperparameters were tuned using GridSearchCV, optimizing the number of weak learners.

The best model achieved an accuracy of 99% on the test set.

Evaluation:

The model's performance was evaluated using the test set.

Precision, recall, and F1-score metrics indicate exceptional performance in classifying breast cancer samples.

## Summary

The Random Forest Regression model achieved a mean squared error of 15.74 on the Boston Housing dataset.

Both the Custom AdaBoost and Sklearn AdaBoost classification models performed well on the Breast Cancer dataset.

The Sklearn AdaBoost model exhibited slightly better performance compared to the Custom AdaBoost model, with higher precision, recall, and accuracy, and fewer misclassifications.

Overall, all models demonstrated effective performance in their respective tasks.

This report summarizes the evaluation of machine learning models on two datasets: the Boston Housing dataset for regression and the Breast Cancer dataset for classification.