Boston housing dataset model step

1. Import all libraries
2. read csv file of boston housing
3. we seet top 5 rows of boston housing
4. we see the shape of data
5. we see the describe the data so we can find min,max,count,standard daviation,25% ,50% and 75% of data
6. we also check null values in data
7. using statemodel we visualize qq plot of medv feature
8. we also visualize distribution plot of medv fearture
9. using heatmap we find correaltion of data we find
    lstat feature and medv feature =highly negative correlated
    dis feature and feature age   =highly negative correlated
    tax feature and feature Rad = =highly positive correlated
10. we start outlier treatment of medv feature
    10.1 we visualize boxplot of medv feature
    10.2 we find 25% and 75% quantile limit and calulate iqr
    10.3 we find upper limit and lower limit using iqr, 25% , 75% qantile limit
    10.4 we see the medv feature data greater than upper limit and lower than lower limit
    10.5 we drop the outlier
11. we splits the data in x & y
12. we feature scaling using minmax scaler
13. we splits the data into train test
14. our first model is lasso regression
    14.1 we train the model we get rmse_score: 7.260585882578913,r2_score: 0.6398735682651674 and accuracy: 63.98735682651674
    14.2 we cross validate the model and we get neg_mean_absolute_error = -5.793866848710175
    14.3 we hyper tune the model using gridsearchcv and cross validate the model we get neg_mean_absolute_error = -5.1911470671652085
15. our second model is ridge regression
    15.1 we train the model we get rmse_score: 6.967377028771655 ,r2_score: 0.6683726894966797 and accuracy: 66.83726894966797
    15.2 we cross validate the model and we get neg_mean_absolute_error = -5.6659347268310984
    15.3 we hyper tune the model using gridsearchcv and cross validate the model we get neg_mean_absolute_error = -5.218625019748393
16. our third model is linear regression we train the model we get rmse_score: 8.671950015277464 ,r2_score: 0.48625785340718586 and accuracy: 48.625785340718586
17. our fourth model is decision tree regression
    17.1 we train the model we get rmse_score: 4.314027507252745 ,r2_score: 0.8728613826815643 and accuracy: 87.28613826815644
    17.2 we cross validate the model and we get neg_mean_absolute_error = -6.717499999999999
    17.3 we hyper tune the model using gridsearchcv and cross validate the model we get neg_mean_absolute_error = -6.9375
18. our fifth model is KNeighbors regression
    18.1 we train the model we get rmse_score: 11.669193059790667 ,r2_score: 0.06976363249939288 and accuracy: 6.976363249939288
    18.2 we cross validate the model and we get neg_mean_absolute_error = -6.328

18.3 we hyper tune the model using gridsearchcv and cross validate the model we get neg_mean_absolute_error = -5.8225

19. our sixth model is Random Forest Regression

19.1 we train the model we get rmse_score: 7.153175681704829 ,r2_score: 0.6504498869018702 and accuracy: 65.04498869018703

19.2 we cross validate the model and we get neg_mean_absolute_error = -5.9919499999999974

19.3 we hyper tune the model using gridsearchcv and cross validate the model we get neg_mean_absolute_error = -5.625500000000001

20. our fifth model is AdaBoost regression

20.1 we train the model we get rmse_score: 4.944941409023705 ,r2_score: 0.8329548139967694 and accuracy: 83.29548139967694

20.2 we cross validate the model and we get neg_mean_absolute_error = -5.40206345529041

20.3 we hyper tune the model using gridsearchcv and cross validate the model we get neg_mean_absolute_error = -5.440608641358642

21. our fifth model is GradientBoosting regression

21.1 we train the model we get rmse_score: 4.908147430689776,r2_score: 0.8354314421814429 and accuracy: 83.54314421814429

21.2 we cross validate the model and we get neg_mean_absolute_error = -5.2364732027329195

21.3 we hyper tune the model using gridsearchcv and cross validate the model we get neg_mean_absolute_error = -5.331728436374019

22. our fifth model is Support Vector regression

22.1 we train the model we get rmse_score: 13.076334778607352 ,r2_score: -0.1681099565672406 and accuracy: -16.81099565672406

22.2 we cross validate the model and we get neg_mean_absolute_error = -5.759032188264307

22.3 we hyper tune the model using gridsearchcv and cross validate the model we get neg_mean_absolute_error = -4.218159543242342

23. we got best model Decision Tree Regression, AdaBoost Regression and GradientBoosting Regression