Regression R2 value

To find out the machine learning regression method using r2 value for the following algorithms and find best parameters.

*Machine Learning:*

The dataset used to find out the r2 value using supervised learning regression method and find the best parameters on the below tables for support vector, decision tree and random forest algorithms. The highlighted R2 value is highest value which was tested and noted here.

1. Multiple Linear regression R2 value is **0.889**
2. Support vector machine regression using below parameters

Parameters:

Kernel:{Linear,Poly,rbf,sigmoid}

C:{10,100,500,1000,2000,3000}

Support Vector machine:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **S.No** | **Tuning value C** | **Linear R2 score** | **Poly**  **R2 score** | **rbf[default]**  **R2score** | **Sigmoid**  **R2 score** |
| 1 | 10 | 0.0018 | -0.004 | -0.005 | -0.005 |
| 2 | 100 | 0.069 | 0.003 | -0.001 | 0.007 |
| 3 | 500 | 0.300 | 0.041 | 0.017 | 0.057 |
| 4 | 1000 | 0.535 | 0.088 | 0.039 | 0.116 |
| 5 | 2000 | 0.758 | 0.169 | 0.079 | 0.207 |
| 6 | 3000 | 0.846 | 0.237 | 0.116 | 0.302 |

Support Vector machine regression R2 value **[kernel=Linear and hyper parameter c=3000]** is 0.84

1. The Decision Tree Regression using below parameters

Parameters:

Criterion:{Squared\_error,friedman\_mse,absolute\_error,poisson}

Max\_features:{sqrt,log2}

Splitter :{best,random}

**Decision Tree:**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **S.No** | **criterion** | **max\_features** | **splitter** | **R2 value** |
| 1. | Squared\_error | sqrt | best | 0.787 |
| 2 | Squared\_error | Log2 | best | 0.474 |
| 3 | Squared\_error | sqrt | Random | 0.055 |
| 4 | Squared\_error | Log2 | Random | 0.366 |
| 5 | Friedman\_mse | sqrt | best | 0.887 |
| 6 | Friedman\_mse | Log2 | best | 0.749 |
| 7 | Friedman\_mse | sqrt | Random | 0.373 |
| 8 | Friedman\_mse | Log2 | Random | 0.326 |
| 9 | Absolute\_error | sqrt | best | 0.405 |
| 10 | Absolute\_error | Log2 | best | 0.649 |
| 11 | Absolute\_error | sqrt | Random | 0.869 |
| 12 | Absolute\_error | Log2 | Random | 0.221 |
| 13 | Poisson | sqrt | best | -0.001 |
| 14 | Poisson | Log2 | best | 0.127 |
| 15 | Poisson | sqrt | Random | -0.246 |
| 16 | Poisson | Log2 | Random | 0.685 |

The Decision tree R2 value [criterion=friedman\_mse,max\_features=sqrt,spliter=best] is 0.88

1. The Random Forest using below parameters

Parameters:

Criterion:{Squared\_error,friedman\_mse,absolute\_error,poisson}

Max\_features:{sqrt,log2,None}

n\_estimator :{10,100}

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| S.No | criterion | n\_estimator | max\_features | R value |
| 1 | Squared\_error | 10 | None | 0.896 |
| 2. | Squared\_error | 100 | None | 0.934 |
| 3. | Squared\_error | 10 | Sqrt | 0.534 |
| 4. | Squared\_error | 100 | Sqrt | 0.720 |
| 5. | Squared\_error | 10 | Log2 | 0.833 |
| 6. | Squared\_error | 100 | Log2 | 0.850 |
| 7. | Absolute\_error | 10 | None | 0.915 |
| 8. | Absolute\_error | 100 | None | 0.921 |
| 9. | Absolute\_error | 10 | Sqrt | 0.791 |
| 10. | Absolute\_error | 100 | Sqrt | 0.753 |
| 11. | Absolute\_error | 10 | Log2 | 0.716 |
| 12. | Absolute\_error | 100 | Log2 | 0.780 |
| 13. | Friedman\_mse | 10 | None | 0.939 |
| 14 | Friedman\_mse | 100 | None | 0.919 |
| 15. | Friedman\_mse | 10 | Sqrt | 0.868 |
| 16. | Friedman\_mse | 100 | Sqrt | 0.763 |
| 17. | Friedman\_mse | 10 | Log2 | 0.720 |
| 18. | Friedman\_mse | 100 | Log2 | 0.756 |
| 19. | poisson | 10 | None | 0.929 |
| 20. | Poisson | 100 | None | 0.920 |
| 21. | Poisson | 10 | Sqrt | 0.576 |
| 22. | Poisson | 100 | Sqrt | 0.826 |
| 23. | Poisson | 10 | Log2 | 0.747 |
| 24. | Poisson | 100 | Log2 | 0.753 |

The Random forest R2 value [criterion=friedman\_mse,max\_features=None,n\_estimator=10] is 0.939

R2 value in each algorithm:

**Support vector machine ------R2 value=0.84**

**Decision Tree---------------------R2 value=0.88**

**Random forest ------------------R2 value=0.93 {Highest R2 value}**

**If the R2 value reaches nearest to 1 then the model perfectly fits the data, as part of our testing found the good model which is nearest to 1 in the Random forest algorithm on above parameters.**