R square calculation for Multiple Linear Regression, SVM, Decision Tree

Dataset – 50_Startups

1.**Multiple Linear Regression** – R^2 value = 0.935

2.**Support Vector Machine**

Parameters:

- C*float, default=1.0*
- kernel{*'linear', 'poly', 'rbf', 'sigmoid', 'precomputed'} or callable, default='rbf'*

Without Standardization:

S.No	***Hyper Parameter C***	***Linear R2 value***	***RBF R2 value***	***Poly R2 value***	***Sigmoid R2 value***
1	1	0.895	-0.057	-0.050	-0.057
2	10	-2.437	-0.055	0.025	-0.057
3	100	-357.079	-0.030	0.465	-0.058
4	1000	-36014.020	0.160	0.640	-0.070

With Standardization:

S.No	***Hyper Parameter C***	***Linear R2 value***	***RBF R2 value***	***Poly R2 value***	***Sigmoid R2 value***
1	1	-0.055	-0.057	-0.057	-0.057
2	10	-0.039	-0.056	-0.053	-0.054
3	100	0.106	-0.050	-0.019	-0.030
4	1000	0.780	0.006	0.266	0.185
5	2000	0.876	0.067	0.481	0.397

This **model** is **not suit** for this **dataset**. Because R^2 is not up to the level.

3.Decision Tree

Parameters:

- criterion{*"squared_error", "friedman_mse", "absolute_error", "poisson"}, default="squared_error"*
- splitter*{"best", "random"}, default="best"*
- max_features*int, float or {"sqrt", "log2"}, default=None*

S.No.	*Criterion*	*Splitter*	*max_features*	*R^2 Value*
1. 1				0.910
2.	*squared_error*	*Best*	*None*	0.920
3.	*squared_error*	*Best*	*Sqrt*	0.402
4.	*squared_error*	*Best*	*Log2*	0.398
5.	*squared_error*	*Random*		0.835
6.	*squared_error*	*Random*	*None*	0.915
7.	*squared_error*	*Random*	*Sqrt*	0.608
8.	*squared_error*	*Random*	*Log2*	0.646
9.	*friedman_mse*	*Best*		0.937
10.	*friedman_mse*	*Best*	*None*	0.892
11.	*friedman_mse*	*Best*	*Sqrt*	0.711
12.	*friedman_mse*	*Best*	*Log2*	0.828
13.	*friedman_mse*	*Random*		0.874
14.	*friedman_mse*	*Random*	*None*	0.926
15.	*friedman_mse*	*Random*	*Sqrt*	0.336
16.	*friedman_mse*	*Random*	*Log2*	-0.433
17.	*absolute_error*	*Best*		0.921
18.	*absolute_error*	*Best*	*None*	0.924
19.	*absolute_error*	*Best*	*Sqrt*	-0.357
20.	*absolute_error*	*Best*	*Log2*	0.897
21.	*absolute_error*	*Random*		0.746
22.	*absolute_error*	*Random*	*None*	0.682
23.	*absolute_error*	*Random*	*Sqrt*	0.101
24.	*absolute_error*	*Random*	*Log2*	0.722
25.	*poisson*	*Best*		0.932
26.	*poisson*	*Best*	*None*	0.929
27.	*poisson*	*Best*	*Sqrt*	0.526
28.	*poisson*	*Best*	*Log2*	0.956
29.	*poisson*	*Random*		0.959
30.	*poisson*	*Random*	*None*	0.749
31.	*poisson*	*Random*	*Sqrt*	0.815
32.	*poisson*	*Random*	*Log2*	0.809

In decision tree, criterion='poisson', splitter='random' gives best R^2 value as 0.959

4. **Random Forest**

Parameters:

- n_estimators*int, default=100*
- criterion*{"squared_error", "absolute_error", "friedman_mse", "poisson"}, default="squared_error"*
- max_features*{"sqrt", "log2", None}, int or float, default=1.0*

S.No.	***n_estimators***	***Criterion***	***max_features***	***R^2 Value***
1	*100*	*squared_error*	*None*	0.946
2	*100*	*squared_error*	*Sqrt*	0.759
3	*100*	*squared_error*	*Log2*	0.759
4	*50*	*squared_error*	*None*	0.944
5	*50*	*squared_error*	*Sqrt*	0.683
6	*50*	*squared_error*	*Log2*	0.683
7	*10*	*squared_error*	*None*	0.925
8	*10*	*squared_error*	*Sqrt*	0.519
9	*10*	*squared_error*	*Log2*	0.519
10	*100*	*friedman_mse*	*None*	0.941
11	*100*	*friedman_mse*	*Sqrt*	0.760
12	*100*	*friedman_mse*	*Log2*	0.760
13	*50*	*friedman_mse*	*None*	0.938
14	*50*	*friedman_mse*	*Sqrt*	0.688
15	*50*	*friedman_mse*	*Log2*	0.688
16	*10*	*friedman_mse*	*None*	0.920
17	*10*	*friedman_mse*	*Sqrt*	0.527
18	*10*	*friedman_mse*	*Log2*	0.527
19	*100*	*absolute_error*	*None*	0.945
20	*100*	*absolute_error*	*Sqrt*	0.785
21	*100*	*absolute_error*	*Log2*	0.785
22	*50*	*absolute_error*	*None*	0.940
23	*50*	*absolute_error*	*Sqrt*	0.722
24	*50*	*absolute_error*	*Log2*	0.722
25	*10*	*absolute_error*	*None*	0.928
26	*10*	*absolute_error*	*Sqrt*	0.721
27	*10*	*absolute_error*	*Log2*	0.721
28	*100*	*poisson*	*None*	0.941
29	*100*	*poisson*	*Sqrt*	0.771
30	*100*	*poisson*	*Log2*	0.771
31	*50*	*poisson*	*None*	0.946
32	*50*	*poisson*	*Sqrt*	0.720
33	*50*	*poisson*	*Log2*	0.720
34	*10*	*poisson*	*None*	0.930
35	*10*	*poisson*	*Sqrt*	0.752
36	*10*	*poisson*	*Log2*	0.752

In Random Forest, 2 model gives best answer for R2 value = 0.946