REGRESSION ASSIGNMENT

1. MULTIPLE LINEAR REGRESSION

The R score value for Multiple Linear regression is 0.7894

2. SUPPORT VECTOR MACHINE

S NO	HYPER PARAMETER	RBF R score	LINEAR R score	POLY R score	SIGMOID R score
1	C10	-0.0181	0.5665	0.1593	0.0730
2	C100	0.3906	0.6359	0.7508	0.5275
3	C1000	0.8283	0.7440	0.8605	0.1437
4	C2000	0.8607	0.7414	0.8601	-2.5840
5	C3000	0.8685	0.7414	0.8600	-6.8261

For Support vector the best value for r score is 0.8685 for [**Kernel RBF and C3000**]

3. DECISION TREE

S NO	CRITERION	SPLITTER	MAX FEATURES	R SCORE
1	Squared_error	best	None	0.7030
2	Squared_error	best	sqrt	0.5973
3	Squared_error	best	auto	0.7109
4	Squared_error	best	Log2	0.7610
5	Squared_error	random	None	0.7082
6	Squared_error	random	sqrt	0.6885
7	Squared_error	random	auto	0.7246
8	Squared_error	random	Log2	0.6675
9	friedman_mse	best	None	0.6911

10	friedman_mse	best	sqrt	0.7338
11	friedman_mse	best	auto	0.6975
12	friedman_mse	best	Log2	0.7374
13	friedman_mse	random	None	0.7069
14	friedman_mse	random	sqrt	0.6335
15	friedman_mse	random	auto	0.7194
16	friedman_mse	random	Log2	0.6646
17	absolute_error	best	None	0.6891
18	absolute_error	best	sqrt	0.6442
19	absolute_error	best	auto	0.6827
20	absolute_error	best	Log2	0.7228
21	absolute_error	random	None	0.7048
22	absolute_error	random	sqrt	0.7453
23	absolute_error	random	auto	0.7100
24	absolute_error	random	Log2	0.7444
25	poisson	best	None	0.7172
26	poisson	best	sqrt	0.6730
27	poisson	best	auto	0.7204
28	poisson	best	Log2	0.6840
29	poisson	random	None	0.6922
30	poisson	random	sqrt	0.6480
31	poisson	random	auto	0.6881
32	poisson	random	Log2	0.7234

For Decision tree the best R score value is 0.7610 [criterion – squared error, splitter -best, max features – log2]

4. RANDOM FOREST

S NO	CRITERION	N ESTIMATORS	MAX FEATURES	R SCORE
1	Squared_error	10	NONE	0.8317
2	Squared_error	50	NONE	0.8508
3	Squared_error	100	NONE	0.8566
4	Squared_error	10	sqrt	0.8551
5	Squared_error	50	sqrt	0.8699
6	Squared_error	100	sqrt	0.8709
7	Squared_error	10	Log2	0.8704
8	Squared_error	50	Log2	0.8648
9	Squared_error	100	Log2	0.8678
10	Absolute_error	10	NONE	0.8427
11	Absolute_error	50	NONE	0.8480
12	Absolute_error	100	NONE	0.8602
13	Absolute_error	10	sqrt	0.8664
14	Absolute_error	50	sqrt	0.8647
15	Absolute_error	100	sqrt	0.8706
16	Absolute_error	10	Log2	0.8421
17	Absolute_error	50	Log2	0.8681
18	Absolute_error	100	Log2	0.8734
19	Friedman_mse	10	NONE	0.8486
20	Friedman_mse	50	NONE	0.8538

21	Friedman_mse	100	NONE	0.8546
22	Friedman_mse	10	sqrt	0.8659
23	Friedman_mse	50	sqrt	0.8654
24	Friedman_mse	100	sqrt	0.8726
25	Friedman_mse	10	Log2	0.8395
26	Friedman_mse	50	Log2	0.8721
27	Friedman_mse	100	Log2	0.8716
28	poisson	10	NONE	0.8609
29	poisson	50	NONE	0.8519
30	poisson	100	NONE	0.8517
31	poisson	10	sqrt	0.8579
32	poisson	50	sqrt	0.8665
33	poisson	100	sqrt	0.8695
34	poisson	10	Log2	0.8511
35	poisson	50	Log2	0.8640
36	poisson	100	Log2	0.8682

For random forest the best R score value is 0.8734 **[criterion – absolute error ,max features – log2, n estimators - 100]**

Final Observation:

When compared to all algorithms in **random forest** only the R score is 0.8734 which is higher than other algorithms