

Machine Learning Algorithms

Identifying 'R2 score' to capture the best model based on the accuracy using hyper tuning parameters

Multiple Linear Regression R2 score = 0.9347

Support Vector Machine :

	r2 score				
S.No	Hyper Parameter	Linear	rbf (nonlinear)	poly (nonlinear)	Sigmoid (nonlinear)
1	C = 1	-0.1572	-0.1597	-0.1593	-0.1594
2	C = 10	-0.1336	-0.1586	-0.1544	-0.1557
3	C = 100	0.0548	-0.1479	-0.1074	-0.1195
4	C = 500	0.6221	-0.1032	0.0781	0.0117
5	C = 1000	0.7849	-0.0627	0.2681	0.1496
6	C = 2000	0.8477	0.0039	0.5264	0.4158
7	C = 3000	0.8959	0.0767	0.6802	0.5961
8	C= 4000	0.8993	0.1433	0.7737	0.6608
9	C= 5000	0.8993	0.1971	0.7976	0.7171

The SVM regression algorithm R2 score (Linear & Hyper parameter (C = 4000)) = 0.8993

Decision Tree :

S.No	Criterion	Splitter	max_features	R2 score
1	squared_error	best	sqrt	0.8567
2	squared_error	best	log2	0.5878
3	squared_error	random	sqrt	0.4606
1	squared_error	random	log2	0.6792
5	friedman_mse	best	sqrt	0.4692
6	friedman_mse	best	log2	0.8645
7	friedman_mse	random	sqrt	0.8757
8	friedman_mse	random	log2	0.9456
7	absolute_error	best	sqrt	0.9358
8	absolute_error	best	log2	0.9548
9	absolute_error	random	sqrt	0.9321
10	absolute_error	random	log2	-0.3445
9	poisson	best	sqrt	0.5469
10	poisson	best	log2	0.1696
11	poisson	random	sqrt	0.2339
12	poisson	random	log2	0.375
13	Defaultparameter	-	-	0.9616

The Decision Tree regression algorithm R2 score (with default parameters) = 0.9616

The Decision Tree regression algorithm R2 score (criterion = 'absolute\_error', splitter = 'best', max\_features = 'log2') = 0.9548

Random Forest:

S.No	n_estimators	Criterion	max_features	R2 score
1	50	squared_error	sqrt	0.8196
2	50	squared_error	log2	0.7864
3	100	squared_error	sqrt	0.8067
1	100	squared_error	log2	0.8302
1	1000	squared_error	log2	0.8109
5	50	friedman_mse	sqrt	0.7967
6	50	friedman_mse	log2	0.7972
7	1000	friedman_mse	sqrt	0.8078
8	1000	friedman_mse	log2	0.8156
7	1000	absolute_error	sqrt	0.9358
8	1000	absolute_error	log2	0.8211
9	2000	absolute_error	sqrt	0.801
10	1000	absolute_error	log2	0.8023
9	1000	poisson	sqrt	0.7607
10	1000	poisson	log2	0.7845

The Random Forest regression algorithm R2 score (n\_estimators = 1000, criterion = 'absolute\_error', max\_features = 'sqrt') = 0.9358

Result : For this dataset "Decision Tree Regression" predicts the better model than MLR, SVM, & RF