

To find the evaluation matrix value using R^2 Method for various Regression algorithms.

Pre requisite: Value between 0-1.

1. MLR: **0.9358**
2. Support Vector Machine:

Sr.No	Hyper Parameter	Linear	Poly	RBF	Sigmoid
1	C=1.0	-0.0556	-0.0574	-0.05741	-0.0572
2	C=10	-0.0396	-0.0536	-0.0568	-0.0547
3	C=100	0.1064	-0.0198	-0.0507	-0.030
4	C=1000	0.7802	0.2661	0.0067	0.1850

Result: The Support Vector Machine algorithm, provides R^2 Value = 0.7802.

SVR (kernel="linear",C=1000)

3. Decision Tree:

Sr.No	Criterion	Max_Features	Splitter	R^2 Value
1	<i>mae(squared_error)</i>	auto	best	0.9497
2	<i>mae(squared_error)</i>	sqrt	best	0.5280
3	<i>mae(squared_error)</i>	log2	best	0.6635
4	<i>mae(squared_error)</i>	auto	random	0.7842
5	<i>mae(squared_error)</i>	sqrt	random	0.5546
6	<i>mae(squared_error)</i>	log2	random	0.6512
7	<i>friedman_mse</i>	auto	best	0.9162
8	<i>friedman_mse</i>	sqrt	best	0.170
9	<i>friedman_mse</i>	log2	best	0.7315
10	<i>friedman_mse</i>	auto	random	0.9189
11	<i>friedman_mse</i>	sqrt	random	-0.4831
12	<i>friedman_mse</i>	log2	random	0.1795
13	Mse	auto	best	0.9102
14	Mse	sqrt	best	0.5129
15	Mse	log2	best	0.6281
16	Mse	auto	random	0.9100
17	Mse	sqrt	random	0.6542
18	Mse	log2	random	-0.0184

Result: The decision tree algorithm, provides R^2 Value = 0.9497

DecisionTreeRegressor(criterion='mae',max_features='auto',splitter='best')

4. Random Forest:

Sr.No	Criterion	Max_Features	n_estimators	R ² Value
1	Mse	Auto	100	0.9339
2	Mse	sqrt	100	0.8434
3	Mse	Log2	100	0.8154
4	Mse	Auto	200	0.9394
5	Mse	sqrt	200	0.7572
6	Mse	Log2	200	0.7904
7	Mae	Auto	100	0.9488
8	Mae	sqrt	100	0.8417
9	Mae	Log2	100	0.8100
10	Mae	Auto	200	0.9392
11	Mae	sqrt	200	0.8217
12	Mae	Log2	200	0.7972
13	<i>friedman_mse</i>	Auto	100	0.9376
14	<i>friedman_mse</i>	sqrt	100	0.8144
15	<i>friedman_mse</i>	Log2	100	0.7360
16	<i>friedman_mse</i>	Auto	200	0.9383
17	<i>friedman_mse</i>	sqrt	200	0.8078
18	<i>friedman_mse</i>	Log2	200	0.8159

Result: The Random Forest algorithm, provides R2 Value = 0.9488

`RandomForestRegressor(criterion='mae',max_features='auto',n_estimators=100)`

Conclusion:

As the above models evaluation matrix value, Decision Tree algorithm provides the highest R² Value (0.9497), so this model has been considered as good model and saved for further deployment phase.

Dataset:



50_Startups.csv