Regression-Algorithm-Insurance

Answers:

1) The problem statement is to identify the insurance charges for the client.

2) Total numner of rows and columns are
   1338 rows × 6 columns.

   Columns are 'age', 'sex', 'bmi', 'children', 'smoker', 'charges'

3) The good model is ‘SLR and MLR’ because it gives the R_score value as 1. And in SVM using kernel ”Linear” as parameter the R_score value is 0.9999

INSURANCE

R_Score value for Decission Tree

S.NO	CRITERION	MAX_FEATURES	SPLITTER	R_SCORE
1	Squarred_error	sqrt	best	0.96208
2	Squarred_error	Log2	best	0.938548
3	Squarred_error	sqrt	random	0.960695
4	Squarred_error	Log2	random	0.965617
5	Friedman_mse	sqrt	best	0.95169
6	Friedman_mse	sqrt	random	0.93097
7	Friedman_mse	Log2	best	0.9870
8	Friedman_mse	Log2	random	0.9554
9	Absolute_error	sqrt	best	0.99010
10	Absolute_error	sqrt	random	0.95845
11	Absolute_error	Log2	best	0.99538
12	Absolute_error	Log2	random	0.92505
13	Poisson	sqrt	best	0.93784
14	Poisson	sqrt	random	0.96208
15	Poisson	Log2	best	0.942230
16	Poisson	Log2	random	0.98396

The highest R_score value is found using criterion for decision 0.99538

R_Score value for SVM

S.No	Parameter(Kernel)	C Value	R_score
1	Linear	0.01	0.999999
2	RBF	0.01	-0.0984
3	Sigmoid	0.001	-0.098698

R_Score value for SLR,MLR,Decission Tree,Random Forest

S.No	Algorithm	R_Score value
1	SLR	1.0
2	MLR	1.0
3	Decission Tree	0.99538
4	Random Forest	0.999823
5	SVM(Linear)	0.9999