

Steps :

- 1 Problem statement: Predict the insurance charges
- 2 Total 6 Columns. Need 5 input and one output. Requirements clear and it is coming under supervised learning. ML Regression
- 3 3 nominal data and 3 ordinal data. Need to use dummies for 3 columns and need to change the type
- 4 Need to create models ML regression, SVM and Decision Tree, choose the best model based on r2_score
- 5 List out the r2_score table
- 6 choose the best model

Multiple Linear Regression**0.78****Support Vector Machine :**

Sl n	Hyperparameter	Linear Value	RBF non lin	Poly	Sigmoid
1	C1	-0.01	-0.008	-0.075	-0.075
2	C10	0.462	-0.03	0.03	0.039
3	C100	0.628	0.319	0.616	0.526
4	C1000	0.764	0.81	0.854	0.212

Decision Tree:

Sl n	Criterion	Splitter	max_features	randomstate	R value
1	squared_error	best	sqrt	1	0.724
2	squared_error	best	log2	1	0.724
3	squared_error	best	none	1	0.713
4	squared_error	random	sqrt	1	0.701
5	squared_error	random	log2	1	0.701
6	squared_error	random	none	1	0.668
7	friedman_mse	best	sqrt	1	0.724
8	friedman_mse	best	log2	1	0.724
9	friedman_mse	best	none	1	0.713
10	friedman_mse	random	sqrt	1	0.701
11	friedman_mse	random	log2	1	0.701
12	friedman_mse	random	none	1	0.668
13	absolute_error	best	sqrt	1	0.563
14	absolute_error	best	log2	1	0.563
15	absolute_error	best	none	1	0.694
16	absolute_error	random	sqrt	1	0.74
17	absolute_error	random	log2	1	0.74
18	absolute_error	random	none	1	0.723
19	poisson	best	sqrt	1	0.716
20	poisson	best	log2	1	0.716

21	poisson	best	none	1	0.686
22	poisson	random	sqrt	1	0.655
23	poisson	random	log2	1	0.655
24	poisson	random	none	1	0.663

Finalized model Decision tree - absolute_error random sqrt

Reason : Even though SVM rvalue is 0.854 the results are abnormal an not in line with the data since the c value crossing 1000, hence Decision Tree with the r score 0.74 selected as a final model