Regression Assignment

A client's requirement is, he wants to predict the insurance charges based on the several parameters .The client has provided the dataset of the same.

As a data scientist , you must develop a model which will predict the insurance charges.

Solution :

I would use a ML to predict the insurance charges

ML-Supervised-Regression

Insurance Prediction List

1.Multiple Linear Regression :

r score=0.7894

2.Support Vector Machine :

S.No	Hyper Parameter C	Linear r score	Poly r score	Rbf r score	Sigmoid r score
1	C=10	-0016	-0.0931	-0.0819	-0.0907
2	C=100	0.5432	-0.0997	-0.1248	-0.1181
3	C=300	0.6123	-0.0919	-0.1258	-0.2304
4	C=500	0.6270	-0.0820	-0.1246	-0.4562
5	C=999	0.6340	-0.0555	-0.1175	-1.1663

3.Decision Tree :

Squared error Best	Squared error Random	Friedman mse Best	Friedman mse Random	Absolute error Best	Absolute error Random	Poisson Best	Poisson Random
0.7039	0.6718	0.6890	0.7063	0.6618	0.7473	0.6665	0.6758

4.Random Forest :

S.No	N Estimators	Max features sqrt r score	Max features log2 r score	Max_features auto r score
1	10	0.8612	0.8544	0.8227
2	100	0.8709	0.8713	0.8540
3	300	0.8717	0.8708	0.8551
4	500	0.8713	0.8728	0.8565
5	999	0.8725	0.8733	0.8549

The final best model for regression using machine learning is :

Random Forest n_estimators=999 , max_features ="log2" R^2 Score =0.8733