

Problem Statement:

Develop a new model for predict health insurance charges based on the age, sex, bmi, childer and smoker parameter.

Dataset details:

The dataset is specifying the information about insurance charges. The dataset has multiple input and one output. This multiple linear regression.

Here are the features:

age	sex	bmi	children	smoker
-----	-----	-----	----------	--------

Independent fields : Age, Sex, BMI, Childern and Is Smoker

Dependent fields : Insurance Charge

Preprocessing Method:

There is a nominal data is available in the dataset, so we need to preprocess using one hot encoding algorithm

Here are the model evaluation and best model is Linear kernel with correction of 3000

No	Hyper Parameter	Linear (r value)	RBF (Non Linear)	Poly (r_value)	SigMod (r value)
1	C10	-0.00161763	-0.0819691	-0.09311616	-0.0907832
2	C100	0.54328182	-0.12480368	-0.09976172	-0.11814555
3	C500	0.627046276	-0.12464161	-0.0820288	-0.45629443
4	C1000	0.634036931	-0.11749092	-0.05550594	-1.66590813
5	C2000	0.689326311	-0.10778764	-0.00270245	-5.61643154
6	C3000	0.759089037	-0.09621285	0.048928964	-12.0190481

Decision Tree Models

No	<i>criterion</i>	Splitter	Max Features	R Value
1	<i>squared_error</i>	<i>best</i>	<i>None</i>	0.707627516
2	<i>squared_error</i>	<i>best</i>	<i>sqrt</i>	0.738117225
3	<i>squared_error</i>	<i>best</i>	<i>log2</i>	0.721713998
4	<i>squared_error</i>	<i>random</i>	<i>None</i>	0.723480813
5	<i>squared_error</i>	<i>random</i>	<i>sqrt</i>	0.728163368
6	<i>squared_error</i>	<i>random</i>	<i>log2</i>	0.726140533
7	<i>friedman_mse</i>	<i>best</i>	<i>None</i>	0.686733982
8	<i>friedman_mse</i>	<i>best</i>	<i>sqrt</i>	0.699954467
9	<i>friedman_mse</i>	<i>best</i>	<i>log2</i>	0.728869594
10	<i>friedman_mse</i>	<i>random</i>	<i>None</i>	0.699923934
11	<i>friedman_mse</i>	<i>random</i>	<i>sqrt</i>	0.681997249
12	<i>friedman_mse</i>	<i>random</i>	<i>log2</i>	0.584977089
13	<i>absolute_error</i>	<i>best</i>	<i>None</i>	0.695402527
14	<i>absolute_error</i>	<i>best</i>	<i>sqrt</i>	0.699429673
15	<i>absolute_error</i>	<i>best</i>	<i>log2</i>	0.698224255
16	<i>absolute_error</i>	<i>random</i>	<i>None</i>	0.716456942
17	<i>absolute_error</i>	<i>random</i>	<i>sqrt</i>	0.676764732
18	<i>absolute_error</i>	<i>random</i>	<i>log2</i>	0.751534952
19	<i>poisson</i>	<i>best</i>	<i>None</i>	0.714667632
20	<i>poisson</i>	<i>best</i>	<i>sqrt</i>	0.714230843
21	<i>poisson</i>	<i>best</i>	<i>log2</i>	0.712018404
22	<i>poisson</i>	<i>random</i>	<i>None</i>	0.685138133
23	<i>poisson</i>	<i>random</i>	<i>sqrt</i>	0.680125837
24	<i>poisson</i>	<i>random</i>	<i>log2</i>	0.646979951