Decision\_Tree r2 value for insurance-pre

|  |  |  |  |
| --- | --- | --- | --- |
| s.no | criterion | splitter | r2value |
| 1 | Squarred\_error | best | 0.69973 |
| 2 | Squarred\_error | random | 0.702953 |
| 3 | absolute\_error | best | 0.67047 |
| 4 | absolute\_error | random | 0.793520 |
| 5 | friedman\_mse | best | 0.695538 |
| 6 | friedman\_mse | random | 0.729809 |
| 7 | Poisson | best | 0.74145 |
| 8 | poisson | random | 0.721594 |

SVMR Nonlinear r2value for insurance\_pre =0.810206

Multilinear r2value for insurance \_ pre=0.7894790

RandomForest r2value for insurance\_pre=0.859112

Report

1.MachineLearing-because input and output data are Numerical present in insurance \_pre Assignment

2.supervised Learning-input and output are all present

3.This assignment data has Numericalvalue –Regression Algorithem

4.Total number of rows=1338

Total number of columns=6

5.yes,I am doing pre-processing method string to number catogerical data –Numerical data ie ,Nominal data-One Hot Encoding

dataset=pd.get\_dummies(dataset,dtype=float,drop\_first=True)

6. Final model is RandomForest it has r2 value0.8509112

So,all the other model has below from RandomForest.so I think insurance\_pre Assignment use best value is RandomForest r2 value 0.8509112.

7.Best model is RandomForest r2 value 0.8509112

8.r\_score value nearby 1 –Good model