Assignment regression:-

1)Problem Statement:-

Want to predict the insurance charges

2)Basic Info of Dataset:-

Total number of rows: 1339

Total number of columns: 6

Total number of Input columns: 5[age, BMI, Sex, Children, Smoker]

Total number of output columns: 1[Charges]

3)Pre-processing method:-

We are using ONE HOT ENCODING Algorithm to convert string to nominal data. Here the column (smoker and sex) is converted into nominal data .

4) For MultiLinear:- 0.789479034986701

For Svm:-

Hyper parameter	linear	poly	sigmoid
C10	0.4624	0.03871	0.03930
C100	0.6288	0.6179	0.52761
C500	0.7631	0.8263	0.44460
C1000	0.7649	0.8566	0.2874
C1500	0.7440	0.8580	-0.0674
C2000	0.74404	0.86055	-0.5939
C3000	0.7414	0.8598	-2.1244

For Decision Tree:-

S.no	Criterion	Splitter	Max Features	R score
1	Squared_error	best	sqrt	0.7274
2	Squared_error	random	sqrt	0.6616
3	Squared_error	best	log2	0.73883
4	Squared_error	random	log2	0.61673
5	Squared_error	best	none	0.69730
6	Squared_error	random	none	0.68396
7	Friedman_mse	best	sqrt	0.69112
8	Friedman_mse	random	sqrt	0.6581439
9	Friedman_mse	best	log2	0.7507853
10	Friedman_mse	random	log2	0.66521
11	Friedman_mse	best	none	0.6914987
12	Friedman_mse	random	none	0.6431856
13	Absolute error	best	sqrt	0.601330
14	Absolute error	random	sqrt	0.72513707
15	Absolute error	best	log2	0.67563654

16	Absolute error	random	log2	0.644588
17	Absolute error	best	none	0.6752
18	Absolute error	random	none	0.744140109
19	poisson	best	sqrt	0.688069
20	poisson	random	sqrt	0.63354
21	poisson	best	log2	0.688394
22	poisson	random	log2	0.695703
23	poisson	best	none	0.71422
24	poisson	random	none	0.69625

For Random Tree:-

S.No	N ESTIMATORS	CRITERION	MAX FEATURES	R SCORE
1	50	Squared_error	sqrt	0.87005
2	100	Squared_error	sqrt	0.871935
3	50	Squared_error	log2	0.86688
4	100	Squared_error	log2	0.8703
5	50	Squared_error	none	0.85559
6	100	Squared_error	none	0.85684
7	50	Friedman_mse	sqrt	0.86676
8	100	Friedman_mse	sqrt	0.8710
9	50	Friedman mse	log2	0.86487
10	100	Friedman_mse	log2	0.8707
11	50	Friedman_mse	none	0.85555
12	100	Friedman_mse	none	0.8543
13	50	Absolute error	sqrt	0.87022
14	100	Absolute error	sqrt	0.8686
15	50	Absolute error	log2	0.87501
16	100	Absolute error	log2	0.871366
17	50	Absolute error	none	0.8601
18	100	Absolute error	none	0.855377
19	50	poisson	sqrt	0.86723
20	100	poisson	sqrt	0.870227
21	50	poisson	log2	0.8698
22	100	poisson	log2	0.872341
23	50	poisson	none	0.85512
24	100	poisson	none	0.854979

5)Best Model:-

From the above mentioned models, the best model is Random Tree Algorithm (with R score = 0.86501, N_estimator = 50, Criterion = Asolute_error, Max Features = log2)