Assignment-Regression Algorithm

1. The problem statement for this task is to develop a predictive model that can accurately estimate insurance charges based on a variety of input parameters.

2. The total number of rows of the dataset is 1338 and the total number of columns is 6. Age, Sex, BMI, Number of Children, Smoker, and Insurance charges are the attributes of this dataset. For the model, Insurance charges are taken as a predictive value or output parameter and the remaining attributes are the input parameters.

3. I'm using One Hot Encoding method to convert the categorical (nominal) data, sex, and smoke into numerical values.

4. Finalized Model:

   Random Forest Regression - Phase 01 - Model Creation
   https://github.com/GauthamOfficial/Regression-Assignment/blob/3c692c0e27c633a6b563bf22a638161754db0fc6/Random%20Forest%20Regression%2014.ipynb

   Random Forest Regression - Phase 02 - Deployment
   https://github.com/GauthamOfficial/Regression-Assignment/blob/3c692c0e27c633a6b563bf22a638161754db0fc6/Random%20Forest%20Regression%2014-Deployment.ipynb

5.

Multiple Linear Regression:

R^2 Value = 0.7894790349867009

Support Vector Machine:

Model No	Kernel	C	R^2 Value
1	rbf	1.0	-0.08338238593619329
2	linear	1.0	-0.010102665316081394
3	poly	1.0	-0.07569965570860893
4	sigmoid	1.0	-0.07542924281107188
5	precomputed	1.0	Error
6	rbf	100.0	0.3200317832050831
7	linear	100.0	0.6288792857320359
8	poly	100.0	0.6179569624059795
9	sigmoid	100.0	0.5276103546510407
10	precomputed	100.0	Error

Decision Tree

Model No	criterion	splitter	max_features	R^2 Value
1	squared_error	best	None	0.6888202917930293
2	friedman_mse	best	None	0.6930511595802755
3	absolute_error	best	None	0.6595624801482052
4	poisson	best	None	0.7262947574251482
5	squared_error	random	None	0.7076936471491737
6	friedman_mse	random	None	0.7409235618631129
7	absolute_error	random	None	0.721308986536229
8	poisson	random	None	0.7138843646303337
9	squared_error	best	sqrt	0.7147324638731477
10	friedman_mse	best	sqrt	0.7137099161878033
11	absolute_error	best	sqrt	0.7527183725983612
12	poisson	best	sqrt	0.6981400803843865
13	squared_error	best	log2	0.7623308762811165
14	friedman_mse	best	log2	0.7336727536831218
15	absolute_error	best	log2	0.6944423162092421
16	poisson	best	log2	0.7342448039310562
17	squared_error	random	sqrt	0.6629550821086372
18	friedman_mse	random	sqrt	0.6634008297435043
19	absolute_error	random	sqrt	0.6828062339385954
20	poisson	random	sqrt	0.7258928225426858
21	squared_error	random	log2	0.726170624549359
22	friedman_mse	random	log2	0.7109339997068576
23	absolute_error	random	log2	0.6437007802414406
24	poisson	random	log2	0.7059916351444611

Random Forest

Model No	criterion	max_features	R^2 Value
1	squared_error	1.0	0.8531315922585214
2	friedman_mse	1.0	0.8508071666090538
3	absolute_error	1.0	0.8538406248188284
4	poisson	1.0	0.8566986790910468
5	squared_error	None	0.8515294717348677
6	friedman_mse	None	0.8512456263551222
7	absolute_error	None	0.8510007135549992
8	poisson	None	0.8537000601236043
9	squared_error	sqrt	0.8713679334104092
10	friedman_mse	sqrt	0.8692688283057971
11	absolute_error	sqrt	0.8694063443299074
12	poisson	sqrt	0.8700740879535566
13	squared_error	log2	0.8695459086759099
14	friedman_mse	log2	0.8724813425661652
15	absolute_error	log2	0.8711320318570183
16	poisson	log2	0.8712117997145099