Assignment-Regression Algorithm

1)Identify your problem statement.

To predict insurance charge based on given criteria.

2) Tell basic info about the dataset (Total number of rows, columns)

No of rows:1338

No of columns:6

Input values:5

Output value:1

3) Mention the pre-processing method if you're doing any (like converting string to number – nominal data)

Must change gender & smoker to nominal data.

4) Develop a good model with r2_score. You can use any machine learning algorithm; you can create many models. Finally, you have to come up with final model.

5.) All the research values (r2_score of the models) should be documented. (You can make tabulation or screenshot of the results.)

S.No	Algorithm	Parameter	R2
1	Multilinear		0.78
2	SVM	linear	0.111
3	SVM	Linear, C=100	0.54
4	SVM	Poly	0.064
5	SVM	Poly,C=100	0.0997
6	SVM	rbf	0.084
7	SVM	rbf,C=10	0.0819
8	SVM	rbf,C=100	-0.124
9	SVM	Sigmoid	0.08994
10	SVM	Sigmoid,C=10	-0.09
11	SVM	Sigmoid,C=100	0.118
12	Decision Tree	Poisson,best	0.735
13	Decision Tree	Poisson,random	0.784
14	Decision Tree	friedman_mse,random	0.704
15	Decision Tree	friedman_mse,best	0.688

16	Decision Tree	absolute_error,random	0.738
17	Decision Tree	absolute_error,best	0.651
18	Decision Tree	squared_error,best	0.695
19	Decision Tree	squared_error,random	0.674
20	Random Forest	squared_error,n_estimator=50, random_state=0	0.854
21	Random Forest	squared_error,n_estimator=100, random_state=0	0.835
22	Random Forest	squared_error,n_estimator=100, random_state=1	0.854
23	Random Forest	squared_error,n_estimator=100, random_state=0	0.853
24	Random Forest	absolute_error,n_estimator=100,random_state=0	0.852
25	Random Forest	absolute_error,n_estimator=10,random_state=0	0.835
26	Random Forest	absolute_error,n_estimator=100,random_state=1	0.852
27	Random Forest	Poisson,n_estimator=100, random_state=0	0.852
28	Random Forest	Poisson,n_estimator=100, random_state=1	0.852

6.) Mention your final model, justify why u have chosen the same.

Will go with Random Forest as that gives the maximum r2 value.

Saved Model: finalized_model_Random_Forest_insurance.sav