Problem Statement :

A client's requirement is, he wants to predict the insurance charges based on the several parameters. The Client has provided the dataset of the same.
As a data scientist, you must develop a model which will predict the insurance charges.
1.) Identify your problem statement
2.) Tell basic info about the dataset (Total number of rows, columns)
3.) Mention the pre-processing method if you're doing any (like converting string to number – 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.)
6.) Mention your final model, justify why u have chosen the same.

To find the following machine learning regression method using r2 value

1. Multiple Linear Regression R^2 value = 0.78

2. Decision tree

S.No	Max features	Splitter	R value
1	auto	best	0.6876
2	auto	random	0.6576
3	sqrt	best	0.7216
4	sqrt	random	0.7165
5	Log2	best	0.7318
6	Log2	random	0.7108
7	auto	best	0.7219
8	auto	random	0.7187
9	sqrt	best	0.7301
10	sqrt	random	0.7109

The decision tree regressionuses R^2 value(log2_best=0.7318)

3.Random Forest Hyperparameter Tuning
Model Configuration Comparison

S.No	n_estimators	max_Depth	min_samples_split	min_samples_leaf	max_features	R^2 value
1	50	None	2	1	auto	0.8498
2	100	None	2	1	auto	0.8512
3	150	None	2	1	auto	0.8524
4	200	None	2	1	auto	0.8519
5	100	10	2	1	auto	0.8445
6	100	20	2	1	auto	0.8501
7	100	None	5	1	auto	0.8489
8	100	None	10	1	auto	0.8467
9	100	None	2	2	auto	0.8495
10	100	None	2	4	auto	0.8478
11	100	None	2	1	sqrt	0.8506
12	100	None	2	1	log2	0.8491
13	150	None	2	1	sqrt	0.8531
14	200	20	2	1	sqrt	0.8518

Best Configuration: Random Forest with n_estimators=150, max_depth=None, min_samples_split=2, min_samples_leaf=1, max_features='sqrt' achieves the highest R^2 value of **0.8531**