ASSIGNMENT-REGRESSION

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.

1.) Problem statement identification

Stage	Name	Result
1	Domain selection	Machine learning *(dealing with numbers)*
2	Learning selection	Supervised learning *(clear requirement, i/p and o/p)*
3	Classification/ Regression	Regression *(o/p is numeric data)*

2.) Basic dataset info

- No_rows=1338
- No_columns=6 (age, sex, bmi, no_children, smoker, charges)

3.) Pre-processing method

- 2 categorical nominal data columns → one_hot_encoding → numerical data

4.) Best model Creation

5.) Research documentation

F:\assignment\insurance_hypertuning_log.pdf

6.) Final model justification

BEST MODEL:

Support Vector Machine Regression		
Criterion	**C**	**r2score**
rbf	10000	0.8779952401449918

Among all the tested models, the SVR *achieved the highest* R^2 *score on the test set* ***(0.877)****, indicating that it explains* ***87%*** *of the variance in the target variable. Therefore, SVR is selected as the best-performing model.*