**Problem Statement or Requirement:**

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**

* As per the requirement we need to predict the insurance charges. So that’s output of the AI.
* We have 3 Stages for this request

Stage 1: Domain selection 🡺**Machine Learning**

Stage 2: Learning selection 🡺 **Supervised Learning**

Stage 3: Types selection under the Supervise Learning 🡺 **Regression**

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

* In this data set we have totally 1338 Rows and 6 no of Columns
* Data set have 2 Categorical columns and 4 Numerical columns

**3.) Mention the pre-processing method if you’re doing any**

**(like converting string to number – nominal data)**

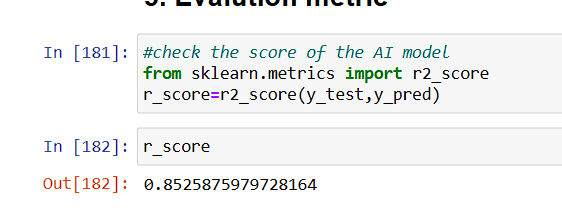
* The dataset mixed with the categorical so we use nominal method which is converting String to number

**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.**

<https://github.com/Mani36Cse/Machine-learning-/tree/main/ML_Regression>

**5.) All the research values (r2\_score of the models) should be documented.**

Doc link : <https://github.com/Mani36Cse/Machine-learning-/tree/main/ML_Regression>



**6.) Mention your final model, justify why u have chosen the same. Kindly create Repository in the name Regression Assignment.**

I choose Random Forest because because large no of dataset so n\_estimators=100, criterion="friedman\_mse" working good comparatively other ML algorithm.