



*Python ML Project 3*

*Linear Regression*

*Insurance Prediction*

### ***Project Description:***

In this particular project, we are using the insurance.csv dataset that contains information like age, sex, bmi, children, smoker, region, charges etc. and using that to predict insurance charges.

However, before you go ahead and make a prediction, it is advised that you first pre-process the data, since it may contain some irregularities and noise.

In addition, try various tricks and techniques in order to gain the best accuracy in your predictions.

### ***Column Details:***

1. **age:** self-explanatory
2. **sex:** male or female
3. **bmi:** body mass index
4. **children:** number of children the person has
5. **smoker:** Yes/No
6. **region:** self-explanatory
7. **charges:** self-explanatory

### ***Part-1: Data Exploration and Pre-processing***

- 1) Load the given dataset
- 2) Fill Null value of children column with the value 0
- 3) Replace the Null values of the column bmi with mean value
- 4) Display a scatter plot between age and children
- 5) Display bar plot between bmi and children
- 6) Perform encoding to convert character data into numerical data
- 7) Perform scaling

### ***Part-2: Working with Models***

- 1) Separate feature data from target data
- 2) Create a Linear regression model between Features and target data
- 3) Display the test score and training score

- 4) Extract slope and intercept value from the model
- 5) Display Mean Squared Error
- 6) Display Mean Absolute Error
- 7) Display Root mean Squared error
- 8) Display R2 score