# Assignment-Regression Algorithm

**Identify your problem statement.**

* AI -----> Prediction
* Stage 1: Domain Selection – Machine Learning.
* Stage 2: Learning Selection – Supervised Learning

Requirement is clear -Predicts the insurance charges.

Dataset i/p & o/p – Clearly mentioned.

* Stage 3: Regression

**Finally ML ---SL----Regression**

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

* Rows – 1338
* Columns – 6 (4 numerical column & 2 categorical column), Input – 5, Output – 1.

**Pre- Processing Methods:**

* Since we have categorical column (Ordinal) , by using Label Encoder mapping has been done to numerical column.

**Model creation & Evaluation metrics:**

* Model has been created by using the below algorithms & respective r2\_score has been published below.

1. Multiple Linear Regression: 0.7894
2. Support Vector Machine:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **S.no** | **Hyper parameter** | ***linear***  **(r2 value)** | **Non-Linear**  rbf  **(r2 value)** | **Non-Linear**  ***poly***  **(r2 value)** | **Non-Linear**  ***sigmoid***  **(r2 value)** |
| 1 | C10 | 0.4624 | -0.0322 | 0.0387 | 0.0393 |
| 2 | C100 | 0.6288 | 0.3200 | 0.6179 | 0.5276 |
| 3 | C500 | 0.7631 | 0.6642 | 0.8263 | 0.4446 |
| 4 | C1000 | 0.7649 | 0.8102 | 0.8566 | 0.2874 |
| 5 | C2000 | 0.7440 | 0.8547 | 0.8605 | -0.5939 |
| 6 | C3000 | 0.7414 | 0.8663 | 0.8598 | -2.1244 |

3. Decision Tree:

|  |  |  |  |
| --- | --- | --- | --- |
| S.no **S.no** | ***criterion*** | ***splitter*** | R2 value |
| 1 | squared\_error | best | 0.6970 |
| 2 | friedman\_mse | best | 0.6880 |
| 3 | absolute\_error | best | 0.6831 |
| 4 | poisson | best | 0.7150 |
| 5 | squared\_error | random | 0.6681 |
| 6 | friedman\_mse | random | 0.7297 |
| 7 | absolute\_error | random | 0.7301 |
| 8 | poisson | random | 0.7186 |

4. Decision Tress:

|  |  |  |
| --- | --- | --- |
| S.no **S.no** | ***criterion*** | R2 value |
| 1 | squared\_error | 0.8537 |
| 2 | friedman\_mse | 0.8539 |
| 3 | absolute\_error | 0.8513 |
| 4 | poisson | 0.8526 |

**Final model:**

**SVM is the best algorithm models among all other algorithm models because of the r2\_score is** 0.8663