**Regression Project**

**Requirement from Client:**

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. **Model Choosing:**

a. The input which is given by the client is in numbers, so we are going to use **Machine Learning**.

1. The Requirement from the Client is very clear, so in here we are going to use **Supervised Learning.**
2. The Label of the project is **Insurance charges (Numeric value)**, so we are going to use **Regression** for this project**.**
3. **Finding the R2-Score for all Regression Algorithm:**

* **Multilinear Regression:**

The value of R2 of Multilinear Regression is 0.7894

* **Support Vector Regression:**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Kernel | | | |
|  | Linear | RBF | Poly | Sigmoid |
| C10 | -0.0016 | -0.0819 | -0.0931 | -0.0907 |
| C100 | 0.5432 | -0.1248 | -0.0997 | -0.1181 |
| C500 | 0.6270 | -0.1246 | -0.0820 | -0.4562 |
| C1000 | 0.6340 | -0.1174 | -0.0555 | -1.6659 |
| C1500 | 0.6394 | -0.1123 | -0.0287 | -3.3163 |
| C2000 | 0.6893 | -0.1077 | -0.0027 | -5.6164 |
| C3000 | 0.7590 | -0.0962 | 0.0489 | -12.019 |

Linear C3000 has the best R2 Score in Support Vector Regression 🡪 0.7590

* **Decision Tree:**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| S.no | Criterion | Splitter | Max Features | R2 Score |
| 1 | Squared Error | Best | Sqrt | 0.6952 |
| 2 | Squared Error | Best | Log2 | 0.7044 |
| 3 | Squared Error | Best | None | 0.7050 |
| 4 | Squared Error | Random | Sqrt | 0.6362 |
| 5 | Squared Error | Random | Log2 | 0.6384 |
| 6 | Squared Error | Random | None | 0.7539 |
| 7 | Friedman\_mse | Best | Sqrt | 0.6937 |
| 8 | Friedman\_mse | Best | Log2 | 0.6817 |
| 9 | Friedman\_mse | Best | None | 0.7123 |
| 10 | Friedman\_mse | Random | Sqrt | 0.6990 |
| 11 | Friedman\_mse | Random | Log2 | 0.6414 |
| 12 | Friedman\_mse | Random | None | 0.7060 |
| 13 | Absolue Error | Best | Sqrt | 0.7168 |
| 14 | Absolue Error | Best | Log2 | 0.6832 |
| 15 | Absolue Error | Best | None | 0.6741 |
| 16 | Absolue Error | Random | Sqrt | 0.6592 |
| 17 | Absolue Error | Random | Log2 | 0.7376 |
| 18 | Absolue Error | Random | None | 0.6882 |
| 19 | Poisson | Best | Sqrt | 0.5167 |
| 20 | Poisson | Best | Log2 | 0.7217 |
| 21 | Poisson | Best | None | 0.7124 |
| 22 | Poisson | Random | Sqrt | 0.6455 |
| 23 | Poisson | Random | Log2 | 0.6768 |
| 24 | Poisson | Random | None | 0.6673 |

Criterion: Squared error, Splitter: Random, Max features: None has the done R2 score in Decision Tree 🡪 0.7539

­­

* **Random Forest:**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| S.No | n\_Estimatore | Criterion | Max features | R2 Score |
| 1 | 50 | Squared Error | Sqrt | 0.8685 |
| 2 | 100 | Squared Error | Sqrt | 0.8714 |
| 3 | 50 | Squared Error | Log2 | 0.8679 |
| 4 | 100 | Squared Error | Log2 | 0.8703 |
| 5 | 50 | Squared Error | None | 0.8555 |
| 6 | 100 | Squared Error | None | 0.8525 |
| 7 | 50 | Friedman\_mse | Sqrt | 0.8664 |
| 8 | 100 | Friedman\_mse | Sqrt | 0.8743 |
| 9 | 50 | Friedman\_mse | Log2 | 0.8691 |
| 10 | 100 | Friedman\_mse | Log2 | 0.8725 |
| 11 | 50 | Friedman\_mse | None | 0.8536 |
| 12 | 100 | Friedman\_mse | None | 0.8509 |
| 13 | 50 | Absolute Error | Sqrt | 0.8723 |
| 14 | 100 | Absolute Error | Sqrt | 0.8674 |
| 15 | 50 | Absolute Error | Log2 | 0.8683 |
| 16 | 100 | Absolute Error | Log2 | 0.8737 |
| 17 | 50 | Absolute Error | None | 0.8555 |
| 18 | 100 | Absolute Error | None | 0.8533 |
| 19 | 50 | Poisson | Sqrt | 0.8708 |
| 20 | 100 | Poisson | Sqrt | 0.8696 |
| 21 | 50 | Poisson | Log2 | 0.8687 |
| 22 | 100 | Poisson | Log2 | 0.8713 |
| 23 | 50 | Poisson | None | 0.8548 |
| 24 | 100 | Poisson | None | 0.8504 |

N\_estimators: 100, Criterion: Friedman\_mse, Max features: Sqrt has the Highest R2 score in Random Forest 🡪 0.8743

1. **Final Model:**
   * By comparing all the Machine Learning Regression Algorithm R2 Score value 🡪 Random Forest (N\_estimators: 100, Criterion: Friedman\_mse, Max features: Sqrt) has the Highest R2 score 🡪 0.8743