Insurance Charges Prediction

**Find following the machine learning regression method using R2 value.**

**1.Multiple Linear Regression (R2-Value = 0.7894)**

**2.Support vector machine**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| S.NO | Hyper  Parameter  (R2Value) | Linear  (R2Value) | RBF (NON-Linear) (R2Value) | Sigmoid  (R2Value) | Poly  (R2Value) |
| **1** | **C10** | **-0.0016** | **-0.0819** | **-0.0907** | **-0.0931** |
| **2** | **C50** | **0.3987** | **-0.1113** | **-0.0985** | **-0.1003** |
| **3** | **C100** | **0.5432** | **-0.1248** | **-0.1181** | **-0.0997** |
| **4** | **C200** | **0.5950** | **-0.1263** | **-0.1575** | **-0.0963** |
| **5** | **C500** | **0.6270** | **-0.1246** | **-0.4562** | **-0.0820** |
| **6** | **C1000** | **0.6340** | **-0.1174** | **-1.6659** | **-0.0555** |
| **7** | **C2000** | **0.6893** | **-0.1077** | **-5.6164** | **-0.0027** |
|  |  |  |  |  |  |
|  |  |  |  |  |  |

The SVM Regression use **R2-Value**(Linear And Hyper Parameter(C2000)) =0.6893

3.Decission Tree

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| S.no | **Criterion** | **Max\_features** | ***Splitter*** | **R2-Value** |  |
| 1 | ***squared\_error*** | **auto** | **best** | **0.7226** |  |
| 2 | ***squared\_error*** | **auto** | **random** | **0.7148** |  |
| 3 | ***squared\_error*** | **sqrt** | **best** | **0.7415** |  |
| 4 | ***squared\_error*** | **sqrt** | **random** | **0.7009** |  |
| 5 | ***squared\_error*** | **Log2** | **best** | **0.7271** |  |
| 6 | ***squared\_error*** | **Log2** | **random** | **0.6541** |  |
| 7 | ***friedman\_mse*** | **auto** | **best** | **0.7075** |  |
| 8 | ***friedman\_mse*** | auto | random | **0.7068** |  |
| 9 | ***friedman\_mse*** | sqrt | best | **0.6799** |  |
| 10 | ***friedman\_mse*** | sqrt | random | 0.7176 |  |
| 11 | ***friedman\_mse*** | Log2 | best | **0.7500** |  |
| 12 | ***friedman\_mse*** | Log2 | random | **0.7067** |  |
| 13 | ***absolute\_error*** | auto | best | **0.6974** |  |
| 14 | ***absolute\_error*** | auto | random | **0.7612** |  |
| 15 | ***absolute\_error*** | sqrt | best | **0.7318** |  |
| 16 | ***absolute\_error*** | sqrt | random | **0.7571** |  |
| 17 | ***absolute\_error*** | Log2 | Best | **0.7497** |  |
| 18 | ***absolute\_error*** | Log2 | random | **0.8110** |  |

The DecisionTree Regression use **R2-Value**(criterion(absolute\_error),max\_features(log2),spliter(random)) =0.8110

4.Randomforest

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| s.no | n\_estimators | criterion | max\_features | **R2-Value** |  |
| 1 | 10 | squared\_error | sqrt | **0.8367** |  |
| 2 | 20 | “ | “ | **0.8508** |  |
| 3 | 30 | “ | “ | **0.8559** |  |
| 4 | 40 | “ | “ | **0.8587** |  |
| 5 | 50 | “ | “ | **0.8603** |  |
| 6 | 60 | “ | “ | **0.8607** |  |
| 7 | 70 | “ | “ | **0.8626** |  |
| 8 | 80 | “ | “ | **0.8630** |  |
| 9 | 90 | “ | “ | **0.8645** |  |
| 10 | 100 | “ | “ | **0.8650** |  |
| 11 | 50 | “ | Log2 | **0.8603** |  |
| 12 | 100 | “ | “ | **0.8650** |  |
| 13 | 50 | absolute\_error | sqrt | **0.8629** |  |
| 14 | 100 | “ | “ | **0.8642** |  |
| 15 | 50 | “ | Log2 | **0.8629** |  |
| 16 | 100 | “ | “ | **0.8642** |  |
| 17 | 50 | ***friedman\_mse*** | sqrt | **0.8602** |  |
| 18 | 100 | “ | “ | **0.8647** |  |
| 19 | 50 | “ | Log2 | **0.8602** |  |
| 20 | 100 | “ | “ | **0.8647** |  |
| 21 | 50 | ***poisson*** | sqrt | **0.8406** |  |
| 22 | 100 | “ | “ | **0.8450** |  |
| 23 | 50 | “ | Log2 | **0.8406** |  |
| 24 | 100 | “ | “ | **0.8450** |  |

The RandomForest Regression use **R2-Value**(n\_estimators=100 criterion(squared\_error),max\_features(log2)) =0.8650