**To find following the Machine Learning Regression method using in $R^2$ value.**

01) **Problem Statement or Requirement is Insurance Charge Prediction.**
02) **Total Number of Rows-6, columns-1338**
03) **I did Preprocessing for sex (male & female) and smoker (yes & no)**

## 1.MULTIPLE LINEAR REGRESSION:

**Best Value from MLR:**

**$R^2$value =** 0.78913454847886

## 2. SUPPORT VECTOR MACHINE:

### i. General Process:

| S.NO | LINEAR (R_VALUE) | RBF(NON LINEAR) (R_VALUE) | POLY (R_VALUE) | SIGMOID (R_VALUE) |
|---|---|---|---|---|
| 01. | -0.111536454002 | -0.088442509991 | -0.064569828857 | -0.089943469577 |

### ii. C_Penalty:

| S.NO | HYPER PARAMETER (C_VALUE) | LINEAR (R_VALUE) | RBF (NON-LINEAR) (R_VALUE) | POLY (R_VALUE) | SIGMOID (R_VALUE) |
|---|---|---|---|---|---|
| 01. | 0.01 | -0.07976404743917187 | -0.08971639718848179 | -0.0893165934318183 | -0.0897333327634073 4 |
| 02. | 0.1 | -0.12212950264630229 | -0.08959641310657296 | -0.0862182772150 4305 | -0.0897653080593388 |
| 03. | 1.0 | -0.11153645400200585 | -0.0884425099 9130221 | -0.0645698288 5737417 | -0.0899434695 7721744 |
| 04. | 10 | -0.00170192 21645011573 | -0.08188099505418855 | -0.0930990339 9514141 | -0.0909307530 0415654 |
| 05. | 100 | 0.543221029625509 | -0.12450697022511292 | -0.0992385759 21484 | -0.1185030051 4767609 |
| 06. | 1000 | 0.6338676426907017 | -0.11761756075675889 | -0.0546562384 41107785 | -1.7112306701 74268 |

### iii. Standardisation:

| S.NO | LINEAR (R_VALUE) | RBF (NON-LINEAR) (R_VALUE) | POLY (R_VALUE) | SIGMOID (R_VALUE) |
|---|---|---|---|---|
| 01. | -0.010195463359872203 | -0.08340516096481387 | -0.07571733771584865 | -0.0754463854853 9218 |

### iv. Standardisation with C_Penalty:

| S.NO | HYPER PARAMETER (C_VALUE) | LINEAR (R_VALUE) | RBF (NON-LINEAR) (R_VALUE) | POLY (R_VALUE) | SIGMOID (R_VALUE) |
|---|---|---|---|---|---|
| 01. | 0.01 | -0.088851825 | -0.089666333 | -0.089589028 | -0.089585755 |
| 02. | 0.1 | -0.080977866 | -0.089095494 | -0.088322830 | -0.088290328 |
| 03. | 1.0 | -0.010195463 | -0.083405160 | -0.075717337 | -0.075446385 |
| 04. | 10 | 0.4624263375 | -0.032380600 | 0.0386251874 | 0.0394401214 |
| 05. | 100 | 0.6289632029 | 0.3196645450 | 0.6164698351 | 0.5268415404 |
| 06. | 1000 | 0.7648394817 | 0.8107195705 | 0.8546515591 | 0.2120454187 |
| 07. | 10000 | 0.7413290358 | 0.8780047074 | 0.8572189614 | -28.34165163 |

**Best Value in SVM:**

**The SVM Regressor $R^2$value (Nonlinear (kernel = "rbf") and Hyper Tuning Parameter (C = 10000)) = 0.8780047074440409**

## 3.DECISION TREE:

### i. Method-01

| SL.NO | CRITERION | SPLITTER | R_VALUE |
|---|---|---|---|
| 01. | Squared error | Best | 0.6883908191578559 |
| 02. | Friedman_mse | Best | 0.6947229954874397 |
| 03. | Absolute error | Best | 0.7256157571187167 |
| 04. | Poisson | Best | 0.681133272801491 |
| 05. | Squared error | Random | 0.7584876928301207 |
| 06. | Friedman_mse | Random | 0.7113158437852314 |
| 07. | Absolute error | Random | 0.7307449040119351 |
| 08. | Poisson | Random | 0.7036416536724851 |

### ii. Method-02

| SL.NO | CRITERION | SPLITTER | MAX_FEATURES | R_VALUE |
|---|---|---|---|---|
| 01. | Squared error | Best | sqrt | 0.7707800713119641 |
| 02. | Squared error | Best | log2 | 0.7614300534063496 |
| 03. | Friedman_mse | Best | sqrt | 0.7445140572039232 |

| 04. | Friedman_mse | Best | log2 | 0.6550595115312327 |
|---|---|---|---|---|
| 05. | Absolute error | Best | sqrt | 0.6796749073935946 |
| 06. | Absolute error | Best | log2 | 0.7114261542959275 |
| 07. | Poisson | Best | sqrt | 0.6472164304876127 |
| 08. | Poisson | Best | log2 | 0.6886508984873159 |
| 09. | Squared error | Random | sqrt | 0.71627952103802 |
| 10. | Squared error | Random | log2 | 0.6799131775438025 |
| 11. | Friedman_mse | Random | sqrt | 0.624486713670842 |
| 12. | Friedman_mse | Random | log2 | 0.6782144931247716 |
| 13. | Absolute error | Random | sqrt | 0.59835828495019 |
| 14. | Absolute error | Random | log2 | 0.6342130151276348 |
| 15. | Poisson | Random | sqrt | 0.6958752685032943 |
| 16. | Poisson | Random | log2 | 0.6003749187441639 |

**Best Value in Decision Tree:**

**The Decision Tree Regressor $R^2$value (criterion='squared_error', splitter='random') =** 0.7707800713119641

## 4.RANDOM FOREST:

### i.Method-1

| S.No | Estimators | R_Value |
|---|---|---|
| 01. | 10 | 0.8254024480527736 |
| 02. | 20 | 0.8366366754276282 |
| 03. | 30 | 0.8464896719139645 |
| 04. | 40 | 0.8513626150908182 |
| 05. | 50 | 0.8519160896863027 |
| 06. | 60 | 0.8504610239826466 |
| 07. | 70 | 0.8513009178708844 |
| 08. | 80 | 0.8514864766488272 |
| 09. | 90 | 0.8519445579703155 |
| 10. | 100 | 0.8524645431942335 |

### ii.Method-2

| S.No | Estimators | CRITERION | R_Value |
|---|---|---|---|
| 01. | 10 | Squared error | 0.8254024480527736 |
| 02. | 50 | Squared error | 0.8519160896863027 |
| 03. | 100 | Squared error | 0.8524645431942335 |
| 04. | 10 | Absolute error | 0.841182394066486 |
| 05. | 50 | Absolute error | 0.8576522515477305 |
| 06. | 100 | Absolute error | 0.8565701539573851 |
| 07. | 10 | Friedman_mse | 0.8254024480527736 |
| 08. | 50 | Friedman_mse | 0.8519160896863027 |
| 09. | 100 | Friedman_mse | 0.8524645431942335 |
| 10. | 10 | Poisson | 0.8280224559126734 |
| 11. | 50 | Poisson | 0.8495669971089854 |
| 12. | 100 | Poisson | 0.8505020311598165 |

### iii.Method-3

| SL.NO | ESTIMATORS | CRITERION | MAX_FEATURES | R_VALUE |
|---|---|---|---|---|
| 01. | 10 | Squared error | sqrt | 0.841371714130463 |
| 02. | 50 | Squared error | sqrt | 0.860384255362958 |
| 03. | 100 | Squared error | sqrt | 0.866171806423152 |
| 04. | 10 | Squared error | log2 | 0.841371714130463 |
| 05. | 50 | Squared error | log2 | 0.860384255362958 |
| 06. | 100 | Squared error | log2 | 0.866171806423152 |
| 07. | 10 | Absolute error | sqrt | 0.844185275693034 |
| 08. | 50 | Absolute error | sqrt | 0.862123099996750 |
| 09. | 100 | Absolute error | sqrt | 0.866661872028384 |
| 10. | 10 | Absolute error | log2 | 0.844185275693034 |
| 11. | 50 | Absolute error | log2 | 0.862123099996750 |
| 12. | 100 | Absolute error | log2 | 0.866661872028384 |
| 13. | 10 | Friedman_mse | sqrt | 0.841371714130463 |
| 14. | 50 | Friedman_mse | sqrt | 0.860384255362958 |
| 15. | 100 | Friedman_mse | sqrt | 0.866171806423152 |

| | | | | |
|---|---|---|---|---|
| 16. | 10 | Friedman_mse | log2 | 0.841371714130463 |
| 17. | 50 | Friedman_mse | log2 | 0.860384255362958 |
| 18. | 100 | Friedman_mse | log2 | 0.866171806423152 |
| 19. | 10 | Poisson | sqrt | 0.836605439397253 |
| 20. | 50 | Poisson | sqrt | 0.854826201808691 |
| 21. | 100 | Poisson | sqrt | 0.859143407373711 |
| 22. | 10 | Poisson | log2 | 0.836605439397253 |
| 23. | 50 | Poisson | log2 | 0.854826201808691 |
| 24. | 100 | Poisson | log2 | 0.859143407373711 |

## Best Value in Random Forest:

**Random Forest Regressor $R^2$value (Estimators = 100, criterion='absolute_error', max_features='sqrt' and 'log2') =** 0.866661872028384

---

## The final Machine Learning Best Algorithm:

**Support Vector Machine output is Best in this Dataset:**

**This is Best model said reason for these model is made the high value of $R^2$value**

**The SVM Regressor $R^2$value (Nonlinear (kernel = "rbf") and Hyper Tuning Parameter (C = 10000)) = 0.8780047074440409**

**(OR)**

**Random Forest Regressor $R^2$value (Estimators = 100, criterion='absolute_error', max_features='sqrt' and 'log2') =** 0.866661872028384