**Insurance Prediction**

1. Clear requirement have both input and output : **Supervised Learning**
2. Input column : **Age, Smoker, sex, bmi,children**
3. Output column: charges
4. Categorical data : Yes
5. It’s Nominal : Yes
6. Algorithm : One-hot-encoding
7. Max rows : 1338
8. Max columns : 6

**ML Metric Value** : 0.7891345484788599

**SVM:**

**Without Hyper ( Not mentioned C value):**

Linear: 0.06025974014685409

RBF (Non-linear): -0.08190532235308945

Poly: --0.06233608549601799

Sigmoid: -0.07205890358470879

**With Hyper Parameter (Mentioned C value)**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **S.No** | **Hyper Parameter** | **Linear** | **RBF** | **Poly** | **Sigmoid** |
| 1 | C = 100 | 0.635766917914623 | 0.39013616263515827 | 0.7508772843937022 | 0.5287870865144102 |
| 3 | C = 1000 | 0.7439340257948153 | 0.8287206345034497 | 0.8591424930442445 | 0.1434938058239389 |
| 4 | C = 2000 | 0.741332614550976 | 0.8596913654584597 | 0.8588677782025483 | -2.5005488010399817 |
| 5 | C = 5000 | 0.7413302488773241 | 0.8729798847256549 | 0.8575581564648183 | -19.837495542575613 |

Decision Tree:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| S.no | Criterion | splitter | Max\_features | metrics |
| 1 | squared\_error | best | None | 0.665595 |
| 2 | Squared\_error | random | None | 0.651923 |
| 3 | Squared\_error | random | sqrt | 0.679469 |
| 4 | Squared\_error | best | Sqrt | 0.611500 |
| 5 | Squared\_error | best | Log2 | 0.631705 |
| 6 | Squared\_error | Random | log2 | 0.697166 |
| 7 | friedman***\_***mse | best | none | 0.702890 |
| 8 | Friedman\_mse | Random | none | 0.728132 |
| 9 | Friedman\_mse | Best | sqrt | 0.748577 |
| 10 | Friedman\_mse | Random | Sqrt | 0.689405 |
| 11 | Friedman\_mse | Best | Log2 | 0.728709 |
| 12 | Friedman\_mse | Random | Log2 | 0.626073 |
| 13 | absolute***\_***error | Best | None | 0.715361 |
| 14 | Absolute\_error | Random | None | 0.689101 |
| 15 | Absolute\_error | Best | sqrt | 0.712328 |
| 16 | Absolute\_error | Random | sqrt | 0.651932 |
| 17 | Absolute\_error | Best | Log2 | 0.729263 |
| 18 | Absolute\_error | Random | Log2 | 0.635313 |
| 19 | poisson | Best | none | 0.669324 |
| 20 | Poisson | Random | none | 0.665361 |
| 21 | Poisson | Best | Sqrt | 0.741746 |
| 22 | Poisson | Random | Sqrt | 0.723797 |
| 23 | Poisson | Best | Log2 | 0.717328 |
| 24 | Poisson | Random | Log2 | 0.6885941 |

Random Forest :

Default values : gini, sqrt 100

Default : 0.851231

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| S.no | Estimator | Criterion | Max\_features | metrics |
| 1 | 100 | gini | Log2 | 0.864581 |
| 2 | 100 | friedman\_mse | Log2 | 0.863773 |
| 3 | 100 | friedman\_mse | sqrt | 0.863840 |
| 4 | 100 | poisson | sqrt | 0.867500 |
| 5 | 100 | poisson | Log2 | 0.864726 |
| 6 | 100 | Squared\_error | Log2 | 0.860059 |
| 7 | 100 | Squared\_error | sqrt | 0.860134 |
| 8 | 100 | Absolute\_error | sqrt | 0.867546 |
| 9 | 100 | Absolute\_error | Log2 | 0.867259 |

Result : For this SVM non-linear Algorithm is the best Model compared to other algorithm.

Value is : 0.8729798847256549