

REGRESSION ASSIGNMENT

1.Domain Selection

Stage 1: Machine Learning

Stage 2: Supervised Learning

Stage 3: Regression

2.Total number of rows=1339

Total number of columns=6

3.By using one hot encoding ,

Dataset=pd.get_dummies(dataset, dtype=int, drop_first=True)

Changed the sex and smokes categorical column to nominal data

5.R2value:

1.Multiple Linear Regression(R2 value)=0.7894

2.Support Vector Machine

| S.No | Hyper parameter | Linear (r2 value) | rbf non-linear (r2 value) | Poly (r2 value) | Sigmoid (r2 value) |
|------|-----------------|-------------------|---------------------------|-----------------|--------------------|
| 1 | C10 | -0.0161 | -0.0322 | 0.0387 | 0.0393 |
| 2 | C100 | 0.5432 | 0.3200 | 0.6179 | 0.5276 |
| 3 | C500 | 0.6270 | 0.6642 | 0.8263 | 0.4446 |
| 4 | C1000 | 0.6340 | 0.8102 | 0.8566 | 0.2874 |
| 5 | C2000 | 0.6893 | 0.8547 | 0.8605 | -0.5939 |
| 6 | C3000 | 0.7590 | 0.8663 | 0.8598 | -2.1244 |

SVM Regression R2 value(rbf non-linear & hyperparameter(C3000)) = 0.8663

3.Decision Tree

| S.No | Criterion | Splitter | R2 Value |
|------|---------------|----------|----------|
| 1 | Squared_error | best | 0.6955 |
| | | random | 0.7246 |
| 2 | Friedman_mse | best | 0.7017 |
| | | random | 0.6776 |

| | | | |
|---|----------------|--------|--------|
| 3 | Absolute_error | best | 0.6593 |
| | | random | 0.7858 |
| 4 | Poisson | best | 0.7347 |
| | | random | 0.7652 |

Decision Tree Regression R2 Value=0.7858

4.Random Forest

| S.No | Criterion | N_Estimators | R2 Value |
|------|----------------|--------------|----------|
| 1 | Squared_error | 10 | 0.8330 |
| | | 50 | 0.8498 |
| | | 100 | 0.8330 |
| 2 | Absolute_error | 10 | 0.8350 |
| | | 50 | 0.8526 |
| | | 100 | 0.8520 |
| 3 | Friedman_mse | 10 | 0.8331 |
| | | 50 | 0.8500 |
| | | 100 | 0.8540 |
| 4 | Poisson | 10 | 0.8313 |
| | | 50 | 0.8491 |
| | | 100 | 0.8526 |

Random Forest R2 Value=0.8540

6.Machine Learning Final model is **SVM Regression** its **R2 Value is 0.8663** than other models its value is near to 1. So I have chosen this model is the best.