

1. Domain- Machine Learning
Learning – Supervised Learning
Regressor or Classification – Regressor
2. Our dataset has 6 columns and 1338 rows
Chargers column is the output column and remaining columns are input columns
3. Data-preprocessing method used for sex and smoker column is ordinal-mapping label encoder method.
4. R2_score for MLR is 0.789
5. SVM

| C | rbf | linear | poly | sigmoid |
|------|---------|--------|--------|---------|
| 1 | -0.083 | -0.010 | -0.075 | -0.075 |
| 10 | -0.0322 | 0.462 | 0.038 | 0.039 |
| 100 | 0.320 | 0.628 | 0.617 | 0.527 |
| 1000 | 0.810 | 0.764 | 0.856 | 0.287 |

Decision Tree

| criterion | splitter | Max_features | R score |
|----------------|----------|--------------|---------|
| Squared_error | Best | Sqrt | 0.752 |
| Friedman_mse | Best | Sqrt | 0.684 |
| Absolute_error | Best | Sqrt | 0.726 |
| Poisson | Best | Sqrt | 0.737 |
| Squared_error | Best | Log2 | 0.766 |
| Friedman_mse | Best | Log2 | 0.735 |
| Absolute_error | Best | Log2 | 0.718 |
| Poisson | Best | Log2 | 0.690 |
| Squared_error | Random | Sqrt | 0.683 |
| Friedman_mse | Random | Sqrt | 0.649 |
| Absolute_error | Random | Sqrt | 0.715 |
| Poisson | Random | Sqrt | 0.605 |
| Squared_error | Random | Log2 | 0.682 |
| Friedman_mse | Random | Log2 | 0.674 |
| Absolute_error | Random | Log2 | 0.789 |
| Poisson | Random | Log2 | 0.638 |
| Squared_error | Best | None | 0.693 |
| Friedman_mse | Best | None | 0.694 |
| Absolute_error | Best | None | 0.658 |
| Poisson | Best | None | 0.720 |

| | | | |
|----------------|--------|------|-------|
| Squared_error | Random | None | 0.709 |
| Friedman_mse | Random | None | 0.700 |
| Absolute_error | Random | None | 0.720 |
| Poisson | Random | None | 0.658 |

Random Forest

| N_estimators | criterion | R score |
|--------------|----------------|---------|
| 100 | Squared_error | 0.849 |
| 100 | Friedman_mse | 0.849 |
| 100 | Absolute_error | 0.848 |
| 100 | Poisson | 0.853 |
| 50 | Squared_error | 0.851 |
| 50 | Friedman_mse | 0.854 |
| 50 | Absolute_error | 0.851 |
| 50 | Poisson | 0.851 |

| N_estimators | Criterion | Max_features | R score |
|--------------|----------------|--------------|---------|
| 100 | Squared_error | Sqrt | 0.872 |
| 100 | Friedman_mse | Sqrt | 0.872 |
| 100 | Absolute_error | Sqrt | 0.873 |
| 100 | Poisson | Sqrt | 0.871 |
| 50 | Squared_error | Sqrt | 0.871 |
| 50 | Friedman_mse | Sqrt | 0.868 |
| 50 | Absolute_error | Sqrt | 0.871 |
| 50 | Poisson | Sqrt | 0.863 |
| 100 | Squared_error | Log2 | 0.871 |
| 100 | Friedman_mse | Log2 | 0.870 |
| 100 | Absolute_error | Log2 | 0.870 |
| 100 | Poisson | Log2 | 0.868 |
| 50 | Squared_error | Log2 | 0.871 |
| 50 | Friedman_mse | Log2 | 0.871 |
| 50 | Absolute_error | Log2 | 0.871 |
| 50 | Poisson | Log2 | 0.868 |

6. Random forest regressor with criterion-‘Absolute_error’ and n_estimators-100 and max_features=‘sqrt’ has the r_score 0.873 with is the highest r_score among all regressors.