

## To predict insurance charges

Domain: Machine Learning

Goal: To predict insurance charges

Learning Selection: Requirement is to predict the insurance charges, and input and output are clear so this comes under supervised learning

Going to predict the insurance charges means going to predict numbers so it comes under ML regression

Project Name: To predict insurance charges

Stage 1: Going to process numbers so Domain is Machine Learning

Stage 2: Supervised Learning

Stage 3: ML Regression

Algorithms to try

1. Multiple Linear

Regression

2. Decision Tree

Regression

3. Support Vector Machine

4. Random Forest

Regression

## Actual Data with Nominal values

]:

|      | age | sex    | bmi    | children | smoker | charges     |
|------|-----|--------|--------|----------|--------|-------------|
| 0    | 19  | female | 27.900 | 0        | yes    | 16884.92400 |
| 1    | 18  | male   | 33.770 | 1        | no     | 1725.55230  |
| 2    | 28  | male   | 33.000 | 3        | no     | 4449.46200  |
| 3    | 33  | male   | 22.705 | 0        | no     | 21984.47061 |
| 4    | 32  | male   | 28.880 | 0        | no     | 3866.85520  |
| ...  | ... | ...    | ...    | ...      | ...    | ...         |
| 1333 | 50  | male   | 30.970 | 3        | no     | 10600.54830 |
| 1334 | 18  | female | 31.920 | 0        | no     | 2205.98080  |
| 1335 | 18  | female | 36.850 | 0        | no     | 1629.83350  |
| 1336 | 21  | female | 25.800 | 0        | no     | 2007.94500  |
| 1337 | 61  | female | 29.070 | 0        | yes    | 29141.36030 |

1338 rows × 6 columns

### Actual Data with numerical values

|      | age | bmi    | children | charges     | sex_male | smoker_yes |
|------|-----|--------|----------|-------------|----------|------------|
| 0    | 19  | 27.900 | 0        | 16884.92400 | 0        | 1          |
| 1    | 18  | 33.770 | 1        | 1725.55230  | 1        | 0          |
| 2    | 28  | 33.000 | 3        | 4449.46200  | 1        | 0          |
| 3    | 33  | 22.705 | 0        | 21984.47061 | 1        | 0          |
| 4    | 32  | 28.880 | 0        | 3866.85520  | 1        | 0          |
| ...  | ... | ...    | ...      | ...         | ...      | ...        |
| 1333 | 50  | 30.970 | 3        | 10600.54830 | 1        | 0          |
| 1334 | 18  | 31.920 | 0        | 2205.98080  | 0        | 0          |
| 1335 | 18  | 36.850 | 0        | 1629.83350  | 0        | 0          |
| 1336 | 21  | 25.800 | 0        | 2007.94500  | 0        | 0          |
| 1337 | 61  | 29.070 | 0        | 29141.36030 | 0        | 1          |

1338 rows × 6 columns

Conclusion:

Best model is Random Forest and r\_score value is 0.883475122

| S. No | n_estimators | random_state | Criterion      | Max features | max_depth | min_samples_split | min_samples_leaf | r score     | Remarks  |
|-------|--------------|--------------|----------------|--------------|-----------|-------------------|------------------|-------------|----------|
| 1     | 200          | 0            | friedman_mse   | sqrt         | 10        | 2                 | 1                | 0.878726296 |          |
| 2     | 200          | 0            | friedman_mse   | log2         | 10        | 2                 | 1                | 0.878726296 |          |
| 3     | 200          | 0            | friedman_mse   | sqrt         | 20        | 2                 | 1                | 0.871251766 |          |
| 4     | 200          | 0            | friedman_mse   | log2         | 20        | 2                 | 1                | 0.871251766 |          |
| 5     | 200          | 0            | friedman_mse   | sqrt         | 30        | 2                 | 1                | 0.871210668 |          |
| 6     | 200          | 0            | friedman_mse   | log2         | 10        | 5                 | 2                | 0.882315637 |          |
| 7     | 200          | 0            | friedman_mse   | sqrt         | 10        | 5                 | 2                | 0.882315637 |          |
| 8     | 200          | 0            | friedman_mse   | sqrt         | 20        | 5                 | 2                | 0.882654802 |          |
| 9     | 200          | 0            | friedman_mse   | log2         | 20        | 5                 | 2                | 0.878726296 |          |
| 10    | 200          | 0            | friedman_mse   | sqrt         | 30        | 5                 | 2                | 0.882654802 |          |
| 11    | 200          | 0            | friedman_mse   | log2         | 30        | 5                 | 2                | 0.882654802 |          |
| 12    | 200          | 0            | friedman_mse   | sqrt         | 10        | 10                | 4                | 0.881828635 |          |
| 13    | 200          | 0            | friedman_mse   | log2         | 10        | 10                | 4                | 0.881828635 |          |
| 14    | 200          | 0            | friedman_mse   | sqrt         | 30        | 10                | 4                | 0.88165423  |          |
| 15    | 200          | 0            | friedman_mse   | log2         | 30        | 10                | 4                | 0.88165423  |          |
| 16    | 200          | 0            | friedman_mse   | log2         | 20        | 10                | 4                | 0.88165423  |          |
| 17    | 200          | 0            | friedman_mse   | sqrt         | 20        | 10                | 4                | 0.88165423  |          |
| 18    | 200          | 0            | friedman_mse   | log2         | 30        | 5                 | 4                | 0.883475122 |          |
| 19    | 200          | 0            | friedman_mse   | sqrt         | 30        | 5                 | 4                | 0.883475122 |          |
| 20    | 200          | 0            | friedman_mse   | Log2         | 20        | 5                 | 4                | 0.883475122 |          |
| 21    | 200          | 0            | friedman_mse   | sqrt         | 20        | 5                 | 4                | 0.883475122 |          |
| 22    | 200          | 0            | friedman_mse   | log2         | 10        | 5                 | 4                | 0.881732361 |          |
| 23    | 200          | 0            | friedman_mse   | sqrt         | 10        | 5                 | 4                | 0.881732361 |          |
| 24    | 100          | 0            | friedman_mse   | sqrt         | 10        | 2                 | 1                | 0.870945766 |          |
| 25    | 100          | 0            | friedman_mse   | log2         | 10        | 2                 | 1                | 0.873526895 |          |
| 26    | 50           | 0            | friedman_mse   | sqrt         | 10        | 2                 | 1                | 0.870044421 |          |
| 27    | 50           | 0            | friedman_mse   | log2         | 10        | 2                 | 1                | 0.870044421 |          |
| 28    |              |              | poisson        |              |           |                   |                  |             | Keyerror |
| 29    |              |              | squared_error  |              |           |                   |                  |             | Keyerror |
| 30    |              |              | absolute_error |              |           |                   |                  |             | Keyerror |