**Gradient Boosting:**

Gradient boosting is a machine learning technique used for building predictive models. Gradient boosting is an ensemble machine-learning technique that combines multiple weak prediction models, to create a powerful predictive model. It sequentially trains models to correct the mistakes of previous models by focusing on the residual errors.

It is effective at capturing complex patterns and is widely used for regression, classification, and ranking tasks.

1. Develop the first model (m1) to predict the mean (average) of the output column.
2. Calculate the “pseudo residual” of the first model (m1) using the formula below

“Pseudo Residual” = “Actual Y” – “Mean”

1. Prepare the second model (m2) to predict the residual. In this model, use the input features (x1,x1…..xn) as inputs, as mentioned instead of predicting Y, it will predict the residual (error).
2. The purpose of the second model (m2) is to assess the intensity of the residual, rather than predicting the Y value directly. To calculate the “Y value”, use the following formula:

Y predict = m1 + m2

* M1 is the output (Mean) of the first model
* m2 is the predicted residual from the second model

1. By using the above formula, you will observe that it predicts the correct “Y value”, which closely matches the **“actual Y value”**. This indicates that the model may be over fitted. To avoid overfitting, introduce a "learning rate" and modify the formula as follows:

Y predict = M1 + (learning rate \* m1)

1. Repeat the above steps until the desired result is achieved.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Sr No | Experience | CGPA | Salary(y) | Base Model  (Mean of output)  (y-pred1) |
| 1 | 90 | 5 | 3 | 4.80 |
| 2 | 100 | 7 | 4 | 4.80 |
| 3 | 110 | 6 | 8 | 4.80 |
| 4 | 120 | 9 | 6 | 4.80 |
| 4 | 80 | 5 | 3 | 4.80 |
| Total | | | |  |

**Step 01:** Develop the first model (m1) to predict the mean (average) of the output column.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |  |  |  | M1 | |
| Sr No | Experience | CGPA | Salary(y) | Base Model  (Mean of output)  (y-pred1) | (y-y-pred =  Residual1) |
| 1 | 90 | 5 | 3 | 4.80 | -1.80 |
| 2 | 100 | 7 | 4 | 4.80 | -0.80 |
| 3 | 110 | 6 | 8 | 4.80 | 3.20 |
| 4 | 120 | 9 | 6 | 4.80 | 1.20 |
| 4 | 80 | 5 | 3 | 4.80 | -1.80 |
| Total | | | |  |  |

**Step 02:** Calculate the “pseudo residual” of the first model (m1) using the formula below

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  | M1 | | M2 |
| Sr No | Experience | CGPA | Salary | Base Model  (Mean of output)  (y-pred1) | (y-y-pred =  Residual1) | Pred1  (Residual1) |
| 1 | 90 | 5 | 3 | 4.80 | -1.80 | -1.80 |
| 2 | 100 | 7 | 4 | 4.80 | -0.80 | -0.80 |
| 3 | 110 | 6 | 8 | 4.80 | 3.20 | 3.20 |
| 4 | 120 | 9 | 6 | 4.80 | 1.20 | 1.20 |
| 5 | 80 | 5 | 3 | 4.80 | -1.80 | -1.80 |
| Total | | | |  |  |  |

**Step 03:** Prepare the second model (m2) to predict the residual. In this model, use the input features (x1,x1…..xn) as inputs,

**Step 04:** then calculate the “Y value(Salary)”, and use the following formula: Y predict = m1 + m2 (where M1 is the output (Mean) of the first model and m2 is the predicted residual from the second model)

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  | M1 | | M2 | |
| Sr No | Experience | CGPA | Salary | Base Model  (Mean of output)  (y-pred1) | (y-y-pred =  Residual1) | Pred1  (Residual1) | Calculate  Salary |
| 1 | 90 | 5 | 3 | 4.80 | -1.80 | -1.80 | 3.00 |
| 2 | 100 | 7 | 4 | 4.80 | -0.80 | -0.80 | 4.00 |
| 3 | 110 | 6 | 8 | 4.80 | 3.20 | 3.20 | 8.00 |
| 4 | 120 | 9 | 6 | 4.80 | 1.20 | 1.20 | 6.00 |
| 4 | 80 | 5 | 3 | 4.80 | -1.80 | -1.80 | 3.00 |
| Total | | | |  |  |  |  |

**Step 05:** By using the above formula, you will observe that it predicts the correct “Y value”, which closely matches the **“actual Y value”**. This indicates that the model may be overfitted. To avoid overfitting, introduce a "learning rate" and modify the formula as follows: Y predict = M1 + (learning rate \* m1)

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  | M1 | | M2 | | |
| Sr No | Experience | CGPA | Salary | Base Model (Mean of output)  (y-pred1) | (y-y-pred = Residual1) | Pred1  (Residual1) | Calculate Salary | Calculate Salary |
| 1 | 90 | 5 | 3 | 4.80 | -1.80 | -1.80 | 3.00 | 4.62 |
| 2 | 100 | 7 | 4 | 4.80 | -0.80 | -0.80 | 4.00 | 4.72 |
| 3 | 110 | 6 | 8 | 4.80 | 3.20 | 3.20 | 8.00 | 5.12 |
| 4 | 120 | 9 | 6 | 4.80 | 1.20 | 1.20 | 6.00 | 4.92 |
| 4 | 80 | 5 | 3 | 4.80 | -1.80 | -1.80 | 3.00 | 4.62 |
| Total | | | |  |  |  |  |  |

**Repeat the above steps until the desired result is achieved (here we have created two more models(M3, M4) are created to get the final result.**

Y predict(m3) = M1 + (learning rate \* m1) +(learning rate \* m2)  
Y predict(m4) = M1 + (learning rate \* m1) +(learning rate \* m2)+ (learning rate \* m3)

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  | M1 | | M2 | | | M3 | | M4 | |
| Sr No | Experience | CGPA | Salary | Base Model (Mean of output) (y-pred1) | (y-y-pred = Residual1) | Pred1 (Residual1) | Calculate Salary | Calculate Salary | Pred1 (Residual1) | Calculate Salary | Pred1 (Residual1) | Calculate Salary |
| 1 | 90 | 5 | 3 | 4.80 | -1.80 | -1.80 | 3.00 | 4.62 | -1.62 | 4.46 | -1.50 | 4.31 |
| 2 | 100 | 7 | 4 | 4.80 | -0.80 | -0.80 | 4.00 | 4.72 | -0.72 | 4.65 | -0.62 | 4.59 |
| 3 | 110 | 6 | 8 | 4.80 | 3.20 | 3.20 | 8.00 | 5.12 | 2.88 | 5.41 | 1.33 | 5.54 |
| 4 | 120 | 9 | 6 | 4.80 | 1.20 | 1.20 | 6.00 | 4.92 | 1.08 | 5.03 | 0.48 | 5.08 |
| 4 | 80 | 5 | 3 | 4.80 | -1.80 | -1.80 | 3.00 | 4.62 | -1.62 | 4.46 | -1.01 | 4.36 |
| Total | | | |  |  |  |  |  |  |  |  |  |