apsara

Date: ___.___.___

Xg-boost (Xtreme Gradient Boosting)

It is a ensemble boosting technique that uses DT as weak learners.

It build tree sequentially, where each new tree is trained to correct the errors of previous trees by minimizing a loss function by gradient descent.

XG Boost improves performance by gradient & hessian along with regularisation making it fast & resistant to overfitting.

Goal → Minimize

↓

Objective → $\sum L(y_i, \hat{y}_i)$ + regularization

⇓

$$\Omega(f) = \gamma T + \frac{1}{2}\lambda \sum w^2$$

T → no. of leaf in tree

γ → It controls depth of tree

λ = L2 regularisation.

Steps of XG Boost

Model starts with one constant prediction for all data points →

Regression → $\hat{y} = \text{mean}(y)$

Classification → $\hat{y} = \log\left(\frac{p}{1-p}\right)$.

apsara

Date: __.__.____

Step 2

Compute loss

Regression → MSE → $L = (y - \hat{y})^2$

Classification → log Loss = $-y \log(p) + (1-y) \log(1-p)$

Step 3

Compute Gradient & hessian (First & Second Derivate)

$$\left\{ g_i = \frac{\partial L}{\partial \hat{y}_i} \right\} \quad \left\{ h_i = \frac{\partial^2 L}{\partial \hat{y}_t^2} \right\}$$

I st order — II nd order

Step 4 Build DT

Split evaluation formula

$$\text{Gain} = \frac{1}{2}\left[\frac{G_L^2}{H_L + \lambda} + \frac{G_R^2}{H_R + \lambda} - \frac{G^2}{H + \lambda} \right] - \gamma$$

$\gamma \rightarrow$ Penalty for split $\quad G = \sum g_i$

$\lambda \rightarrow$ regularisation $\quad H = \sum H_i$

Step 5 Compute leaf weight

$$\left\{ w = - \frac{\sum g_i}{\sum h_i + \lambda} \right\}$$

Step 6 Update Prediction

$$\hat{y}(t) = \hat{y}(t-1) + \eta \cdot f_t(x)$$

Step 7 → Repeat

Final model → $\hat{y}(x) = \sum_{t=1}^{T} \eta f_t(x)$

Final prediction → sum of all tree corelations.