

Gradient Boosting

Gradient Boosting Model

$$\mathbf{F}_{n+1}(\mathbf{X}) = \mathbf{F}_n(\mathbf{X}) + \gamma_n \mathbf{H}(\mathbf{x}, \mathbf{e}_n)$$

$$\begin{aligned}
 \mathbf{F}_0(\mathbf{X}) &= \text{?} \text{?} \mathbf{H}_0(\mathbf{x}, \mathbf{y}) + \mathbf{e}_0 \\
 \mathbf{F}_1(\mathbf{X}) &= \text{?}^0 \mathbf{F}_0(\mathbf{X}) + \text{?} \text{?} \mathbf{H}_1(\mathbf{x}, \mathbf{e}_0) + \mathbf{e}_1 \\
 \mathbf{F}_2(\mathbf{X}) &= \text{?}^1 \mathbf{F}_1(\mathbf{X}) + \text{?} \text{?} \mathbf{H}_2(\mathbf{x}, \mathbf{e}_1) + \mathbf{e}_2 \\
 &\vdots \\
 \mathbf{F}_n(\mathbf{X}) &= \text{?}^n \mathbf{F}_{n-1}(\mathbf{X}) + \text{?} \text{?} \mathbf{H}_n(\mathbf{x}, \mathbf{e}_{n-1}) + \mathbf{e}_n
 \end{aligned}$$

Gradient Boosting Model

$$\mathbf{F}_{n+1}(\mathbf{X}) = \mathbf{F}_n(\mathbf{X}) + \gamma_n \mathbf{H}(\mathbf{x}, \mathbf{e}_n)$$

$$\begin{aligned}
 \mathbf{F}_0(\mathbf{X}) &= \text{?} \mathbf{H}_0(\mathbf{x}, y) + \mathbf{e}_0 \\
 \mathbf{F}_1(\mathbf{X}) &= \mathbf{F}_0(\mathbf{X}) + \text{?} \mathbf{H}_1(\mathbf{x}, \mathbf{e}_0) + \mathbf{e}_1 \\
 \mathbf{F}_2(\mathbf{X}) &= \mathbf{F}_1(\mathbf{X}) + \text{?} \mathbf{H}_2(\mathbf{x}, \mathbf{e}_1) + \mathbf{e}_2 \\
 &\vdots \\
 \mathbf{F}_n(\mathbf{X}) &= \mathbf{F}_{n-1}(\mathbf{X}) + \text{?} \mathbf{H}_n(\mathbf{x}, \mathbf{e}_{n-1}) + \mathbf{e}_n
 \end{aligned}$$

Gradient Boosting Model

$$\mathbf{F}_{n+1}(\mathbf{X}) = \mathbf{F}_n(\mathbf{X}) + \gamma_n \mathbf{H}(\mathbf{x}, \mathbf{e}_n)$$

$$L = (y - y')^2$$

Gradient Boosting Model

$$\mathbf{F}_{n+1}(\mathbf{X}) = \mathbf{F}_n(\mathbf{X}) + \gamma_n \mathbf{H}(\mathbf{x}, \mathbf{e}_n)$$

$$L = (y - y')^2$$

$$L = (y - F_n(x))^2$$

Gradient Boosting Model

$$\mathbf{F}_{n+1}(\mathbf{X}) = \mathbf{F}_n(\mathbf{X}) + \gamma_n \mathbf{H}(\mathbf{x}, \mathbf{e}_n)$$

$$L = (y - y')^2$$

$$dL = (y - F_n(x))^2$$

$$\frac{dL}{dF_n(x)} = 2(y - F_n(x))$$

$$dF_n(x)$$

Gradient Boosting Model

$$\mathbf{F}_{n+1}(\mathbf{X}) = \mathbf{F}_n(\mathbf{X}) + \gamma_n \mathbf{H}(\mathbf{x}, \mathbf{e}_n)$$

$$L = (y - y')^2$$

$$dL = (y - F_n(x))^2$$

$$-\frac{dL}{dF_n(x)} = 2(y - F_n(x))$$

$$dF_n(x)$$

Gradient Boosting Model

$$\mathbf{F}_{n+1}(\mathbf{X}) = \mathbf{F}_n(\mathbf{X}) + \gamma_n \mathbf{H}(\mathbf{x}, \mathbf{e}_n)$$

$$L = (y - y')^2$$

$$dL = (y - F_n(x))^2$$

$$-\frac{2(y - F_n(x))}{dF_n(x)}$$

$$dF_n(x)$$

Gradient Boosting Model

$$\mathbf{F}_{n+1}(\mathbf{X}) = \mathbf{F}_n(\mathbf{X}) + \gamma_n \mathbf{H}\left(\mathbf{x}, -\frac{dL}{dF_n(\mathbf{x})}\right)$$

$$L = (y - y')^2$$

$$dL = (y - F_n(x))^2$$

$$-\frac{dL}{dF_n(x)} = 2(y - F_n(x))$$

$$dF_n(x)$$

Gradient Boosting Model

$$\mathbf{F}_{n+1}(\mathbf{X}) = \mathbf{F}_n(\mathbf{X}) + \gamma_n \mathbf{H}\left(\mathbf{x}, -\frac{dL}{dF_n(\mathbf{x})}\right)$$

Gradient Boosting Model

$$\mathbf{F}_{n+1}(\mathbf{X}) = \mathbf{F}_n(\mathbf{X}) + \gamma_n \mathbf{H}\left(\mathbf{x}, -\frac{dL}{dF_n(\mathbf{x})}\right)$$

↓

$$\text{Loss} = L(y, F_n(x))$$

Gradient Boosting Model

$$\mathbf{F}_{n+1}(\mathbf{X}) = \mathbf{F}_n(\mathbf{X}) + \gamma_n \mathbf{H}\left(\mathbf{x}, -\frac{dL}{dF_n(\mathbf{x})}\right)$$

$$\text{Loss} = L(y, F_n(x)) + \gamma_n L\left(-\frac{dL}{dF_n(x)}, -\right)$$

Gradient Boosting Model

$$\mathbf{F}_{n+1}(\mathbf{X}) = \mathbf{F}_n(\mathbf{X}) + \gamma_n \mathbf{H}\left(\mathbf{x}, -\frac{dL}{dF_n(\mathbf{x})}\right)$$

$$\text{Loss} = \underbrace{L(y, F_n(x))}_{\text{Loss}} + \gamma_n L\left(-\frac{dL}{dF_n(x)}, -\right)$$

Gradient Boosting Model

$$\mathbf{F}_{n+1}(\mathbf{X}) = \mathbf{F}_n(\mathbf{X}) + \gamma_n \mathbf{H}\left(\mathbf{x}, -\frac{dL}{dF_n(\mathbf{x})}\right)$$

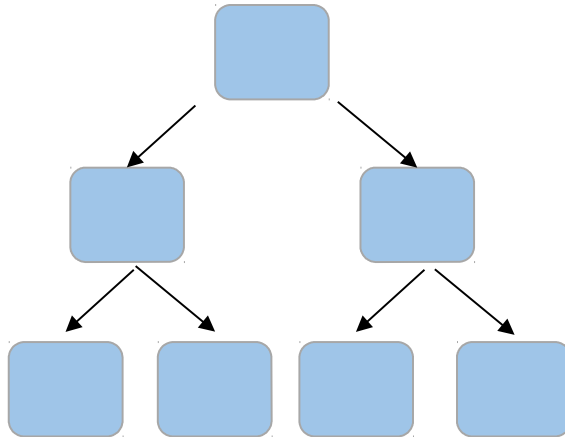
$$\text{Loss} = L(y, F_n(x)) + \underbrace{\gamma_n L\left(-\frac{dL}{dF_n(x)}, -\right)}_{\text{ }}$$

Gradient Boosting Decision Tree

$$\mathbf{F}_{n+1}(\mathbf{X}) = \mathbf{F}_n(\mathbf{X}) + \gamma_n \mathbf{H}\left(\mathbf{x}, -\frac{dL}{dF_n(\mathbf{x})}\right)$$

Gradient Boosting Decision Tree

$$\mathbf{F}_{n+1}(\mathbf{X}) = \mathbf{F}_n(\mathbf{X}) + \gamma_n \mathbf{H}\left(\mathbf{x}, -\frac{dL}{dF_n(\mathbf{x})}\right)$$



Gradient Boosting Decision Tree

$$\mathbf{F}_{n+1}(\mathbf{X}) = \mathbf{F}_n(\mathbf{X}) + \gamma_n \mathbf{H}\left(\mathbf{x}, -\frac{dL}{dF_n(\mathbf{x})}\right)$$

