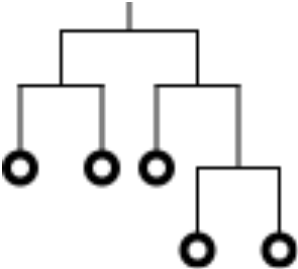


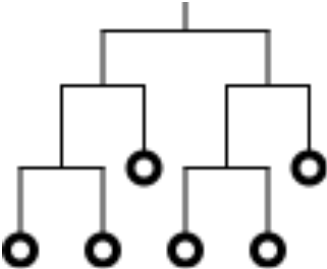


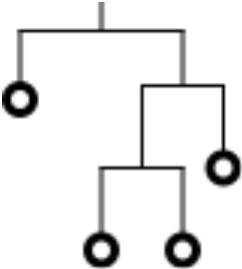




Gradient Boosting Machine (GBM)

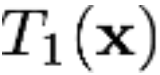












I 2



13





T

nn

(X)



---

**Algorithm 1** Forward Stagewise Additive Training

---

- 1: Initialize  $f_0(\mathbf{x}) = 0$ .
  - 2: **for**  $i = 1$  to  $M$  **do**
  - 3:      $\hat{\Theta}_i = \arg \min_{\Theta_i} \sum_{i=1}^n L(y_i, f_{i-1}(x_i) + T(x_i; \Theta_i))$
  - 4:      $f_i(\mathbf{x}) = f_{i-1} + T(\mathbf{x}; \hat{\Theta}_i)$
  - 5: **end for**
  - 6:  $f(\mathbf{x}) = f_M(\mathbf{x})$
-

# GBM Algorithm

---

**Algorithm 10.3** *Gradient Tree Boosting Algorithm.*

---

1. Initialize  $f_0(x) = \arg \min_{\gamma} \sum_{i=1}^N L(y_i, \gamma)$ .

2. For  $m = 1$  to  $M$ :

(a) For  $i = 1, 2, \dots, N$  compute

$$r_{im} = - \left[ \frac{\partial L(y_i, f(x_i))}{\partial f(x_i)} \right]_{f=f_{m-1}}.$$

(b) Fit a regression tree to the targets  $r_{im}$  giving terminal regions  $R_{jm}$ ,  $j = 1, 2, \dots, J_m$ .

(c) For  $j = 1, 2, \dots, J_m$  compute

$$\gamma_{jm} = \arg \min_{\gamma} \sum_{x_i \in R_{jm}} L(y_i, f_{m-1}(x_i) + \gamma).$$

(d) Update  $f_m(x) = f_{m-1}(x) + \sum_{j=1}^{J_m} \gamma_{jm} I(x \in R_{jm})$ .

3. Output  $\hat{f}(x) = f_M(x)$ .

---

Springer Series in Statistics

Trevor Hastie  
Robert Tibshirani  
Jerome Friedman

## The Elements of Statistical Learning

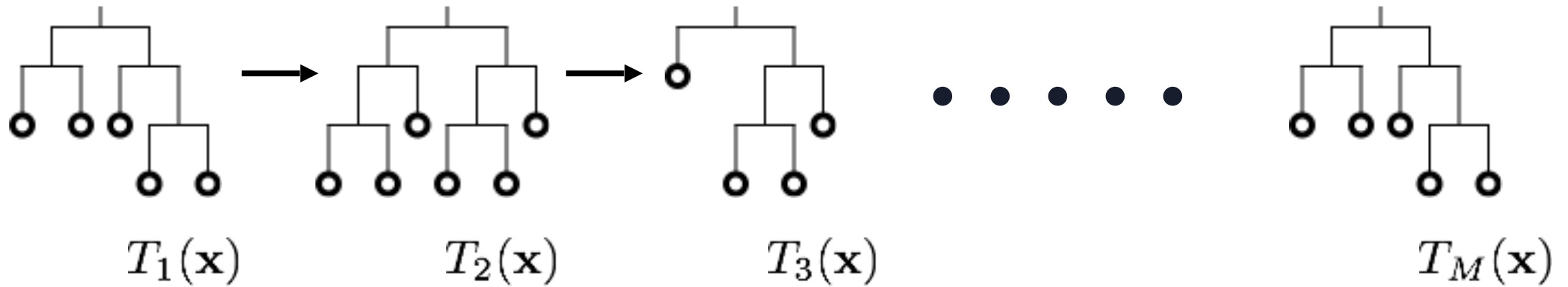
Data Mining, Inference, and Prediction

Second Edition

 Springer



# Gradient Boosting Machine (GBM)



---

## Algorithm 1 Forward Stagewise Additive Training

---

- 1: Initialize  $f_0(\mathbf{x}) = 0$ .
  - 2: **for**  $i = 1$  to  $M$  **do**
  - 3:    $\hat{\Theta}_i = \arg \min_{\Theta_i} \sum_{i=1}^n L(y_i, f_{i-1}(x_i) + T(x_i; \Theta_i))$
  - 4:    $f_i(\mathbf{x}) = f_{i-1} + T(\mathbf{x}; \hat{\Theta}_i)$
  - 5: **end for**
  - 6:  $f(\mathbf{x}) = f_M(\mathbf{x})$
-