

Part.04

Ensemble Learning

I Gradient Boosting

FASTCAMPUS
ONLINE

머신러닝과 데이터분석 A-Z

강사. 이경택

I Gradient Boosting

■ Boosting이란

- Boosting은 오분류된 데이터에 초점을 맞추어 더 많은 가중치를 주는 방식
- 초기에는 모든 데이터가 동일한 가중치를 가지지만, 각 round가 종료된 후 가중치와 중요도를 계산
- 복원추출 시에 가중치 분포를 고려
- 오분류된 데이터가 가중치를 더 얻게 됨에 따라 다음 round에서 더 많이 고려됨
- Boosting 기법으로 AdaBoost, LPBoost, TotalBoost, BrownBoost, MadaBoost, LogitBoost, Gradient Boosting 등이 있음

I Gradient Boosting

■ Gradient Boosting

- Boosting 기법들 간 차이는 오분류된 데이터를 다음 Round에서 어떻게 반영할 것인가의 차이
- AdaBoost는 오분류된 데이터들에 더 큰 가중치를 주어 다음 Round 샘플링에 반영
- Gradient Boosting은 이전 Round의 합성 분류기 $H_{t-1}(x) = \sum_{m=1}^{t-1} \alpha_m h_m(x)$ 의 데이터 별 오류를 예측 하는 새로운 약한 분류기를 학습

I Gradient Boosting

■ Gradient Boosting

- x 를 입력 받아 y 를 예측하는 모델 h_0 가 있다고 하자.

$$y = h_0(x) + \text{error}$$

- Error가 예측 불가능한 랜덤 노이즈가 아닌 경우, 예측 성능을 올리는 가장 직관적인 방법은 error를 제거하는 것.
- 그렇다면 어떻게 error를 제거할 수 있을까?

$$\text{error} = h_1(x) + \text{error2}$$

$$\text{error2} = h_2(x) + \text{error3}$$

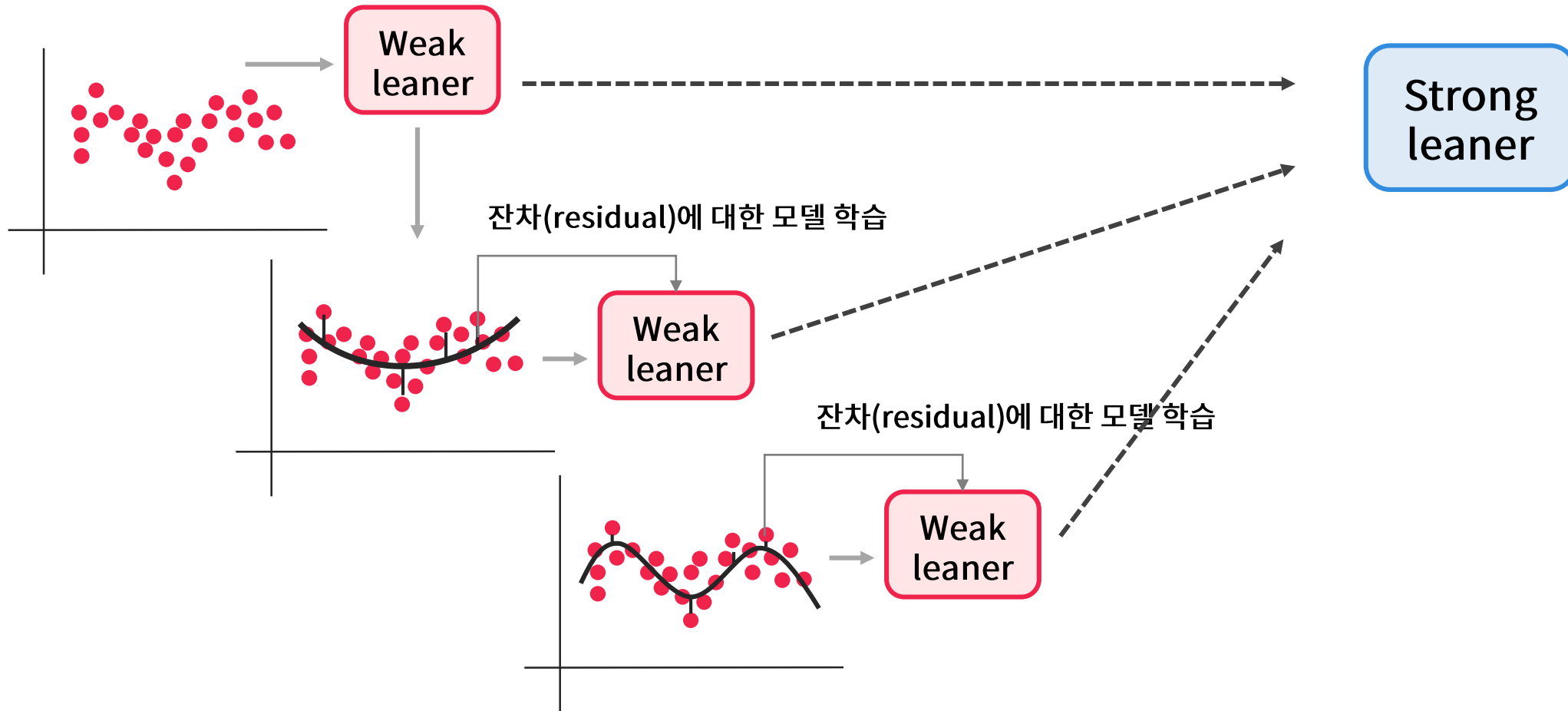
⋮



$$y = h_0(x) + h_1(x) + h_2(x) + \cdots + \text{small error}$$

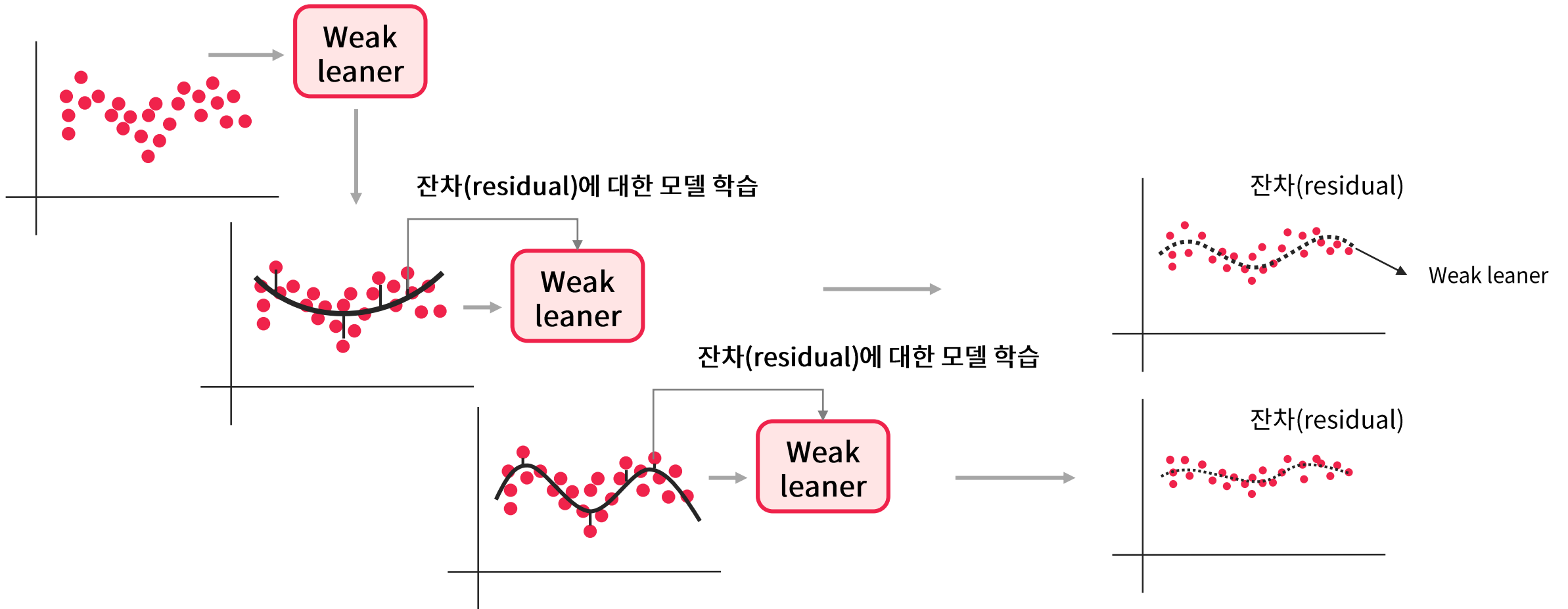
I Gradient Boosting

■ Gradient Boosting



I Gradient Boosting

■ Gradient Boosting



I Gradient Boosting

■ Gradient란?

- $Loss$ 함수를 다음과 같이 정의

$$L(y_i, f(x_i)) = \frac{1}{2} (y_i - f(x_i))^2$$

- 이때의 $Loss$ 함수의 기울기(gradient)

$$\frac{\partial L(y_i, f(x_i))}{\partial f(x_i)} = \frac{\partial [\frac{1}{2} (y_i - f(x_i))^2]}{\partial f(x_i)} = f(x_i) - y_i$$

- Negative gradient = Residual

$$-(f(x_i) - y_i) = y_i - f(x_i)$$

- Negative gradient(residual)를 최소화 시키면서 학습 시키기때문에 gradient boosting이라 부름

Part.04

Ensemble Learning

| Gradient Boosting의 종류

FASTCAMPUS
ONLINE

머신러닝과 데이터분석 A-Z

강사. 이경택