

# AdaBoost

핵심 한 줄

약한 모델을 순차적으로 결합해, 틀린 샘플에 더 집중하도록 가중치를 조정하며 성능을 끌어올리는 양상을

주로 쓰는 곳

이진/다중 분류, 간단한 베이스 학습기(결정 스템프)로 빠르게 베이스라인 만들 때

## AdaBoost (Adaptive Boosting)

약한 학습기를 순차적으로 결합해 “강한” 분류기를 만드는 양상을

### 한 줄 요약

- 틀린 샘플에 더 큰 가중치를 주며 학습을 반복
- 약한 학습기 (주로 결정 스템프)를 가중 합으로 결합
- 작은 모델로도 의외로 강력한 성능을 내는 고전적인 부스팅



그림 1. AdaBoost 개념 요약(약한 학습기의 순차적 결합)

# 1. AdaBoost란?

AdaBoost(Adaptive Boosting)는 **부스팅(Boosting)** 계열의 대표 알고리즘으로, 약한 학습기(weak learner)를 여러 번 학습시키되 **이전 단계에서 틀린 샘플에 더 큰 가중치**를 부여해 다음 학습기가 그 샘플을 더 잘 맞추도록 유도합니다.

- 기본 아이디어: '어려운(자주 틀리는) 데이터'에 점점 더 집중한다.
- 베이스 학습기: 보통 결정 스템프(깊이 1 의사결정트리)처럼 매우 단순한 모델을 사용한다.
- 최종 예측: 각 학습기의 성능에 비례한 가중치( $\alpha$ )를 두고 가중 합(분류는 가중 다수결)을 낸다.



그림 2. 부스팅 과정 시각화 예시(오분류 샘플 강조 → 다음 학습기 보완)

## 2. 학습 절차(핵심 로직)

### 단계

### 설명

#### (1) 샘플 가중치 초기화

모든 샘플 가중치  $w_i$  를 동일하게 시작한다.

#### (2) 약한 학습기 학습

현재 가중치 분포  $w$  로 학습기를 한 번 학습한다.

#### (3) 오차율 계산

가중 오차  $\epsilon = \sum w_i \cdot [y_i \neq h(x_i)]$  를 계산한다.

#### (4) 학습기 가중치 결정

$\alpha = 0.5 \cdot \ln((1-\epsilon)/\epsilon)$ . (오차가 작을수록  $\alpha$  가 커져 영향력이 커진다.)

#### (5) 샘플 가중치 업데이트

틀린 샘플의 가중치는 키우고, 맞춘 샘플의 가중치는 줄인다. 예:  $w_i \leftarrow w_i \cdot \exp(\pm \alpha)$  후 정규화

### 3. 장단점과 실무 팁

장점	주의점/단점
----	--------

- 단순 베이스 학습기만으로도 성능 향상이 빠르다.
- 특성 스케일링에 덜 민감한 편(트리 기반 베이스일 때).
- 해석: 각 약한 학습기의 영향( $\alpha$ )을 통해 어느 정도 설명 가능.
- 라벨 노이즈/이상치에 취약(틀린 샘플을 끝까지 크게 학습하려는 경향).
- 학습기는 순차 학습이라 병렬화가 제한적.
- 학습기 수가 커지면 과적합/불안정 가능(데이터/파라미터에 따라).

**실무 팁:** 결정 스텁프(depth=1) + n\_estimators(학습기 개수) + learning\_rate 를 먼저 튜닝하고, 노이즈가 많다면(라벨 오류/이상치) 로버스트한 모델(예: Gradient Boosting 계열)도 함께 비교해보는 게 좋다.

### 4. 변형/관련 개념

- **SAMME / SAMME.R:** 다중 분류용 AdaBoost 확장(라이브러리에서 옵션으로 제공되는 경우가 많음).
- **AdaBoostRegressor:** 회귀 버전(오차 기반으로 샘플 가중치를 조정).
- **Gradient Boosting:** AdaBoost 의 '순차 보완' 아이디어를 일반화해, 손실 함수의 기울기(gradient)를 따라 약한 학습기를 추가하는 계열.



그림 3. AdaBoost 학습 흐름 예시(가중치 업데이트와 약한 학습기 결합)

## 5. 요약

- AdaBoost 는 '틀린 샘플을 더 크게 보자'라는 직관을 수학적으로 구현한 부스팅 알고리즘이다.
- 약한 학습기들을 순서대로 학습시키며, 각 학습기는 이전 단계의 실수를 보완한다.
- 노이즈/이상치가 많을수록 성능이 불안정해질 수 있으므로 데이터 품질과 규제(learning\_rate, 학습기 수)를 함께 본다.

참고(대표 구현): scikit-learn 의 AdaBoostClassifier / AdaBoostRegressor 문서 및 예제.