



Boosting

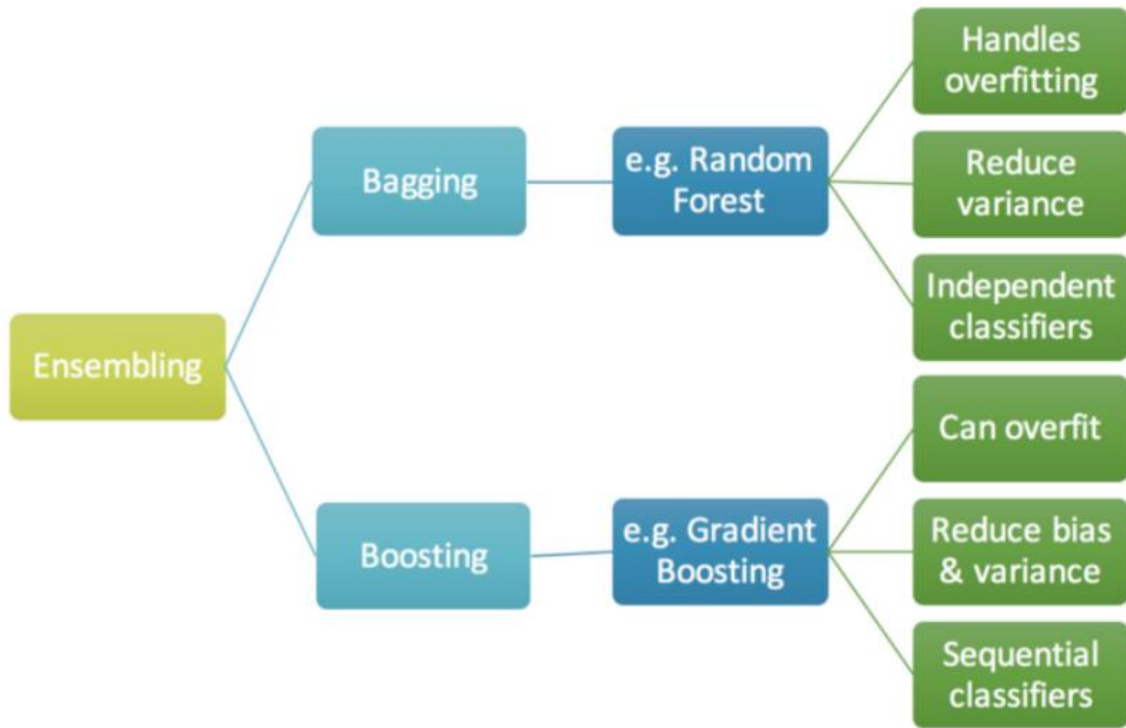
고유경



Index

1. Bagging vs. Boosting
2. Ada Boosting
3. Gradient Boosting
4. XG Boosting

Bagging vs. Boosting



- 병렬적 모델결합
- 독립적으로 모델 구성
- 매 sampling마다 동일 가중치 부여

- 직렬적 모델결합
- 이전 모델의 오류를 바탕으로 새 모델 구성
- 학습오류 큰 데이터에 가중치 부여
- 단일 모델의 성능 낮을 경우

Boosting

간단하고 성능이 떨어지는 **약분류기(weak classifier)**들이
상호보완하도록 단계적으로 학습하고,
이를 조합하여 만들어진 **최종 강분류기(strong classifier)**의
성능을 증폭시키는 원리

Ex) 분류기 A,B,C의 accuracy가 각각 0.3일 때,
이를 결합하여 더 높은 정확도(ex. 0.7)을 얻는 것

1. Ada Boosting

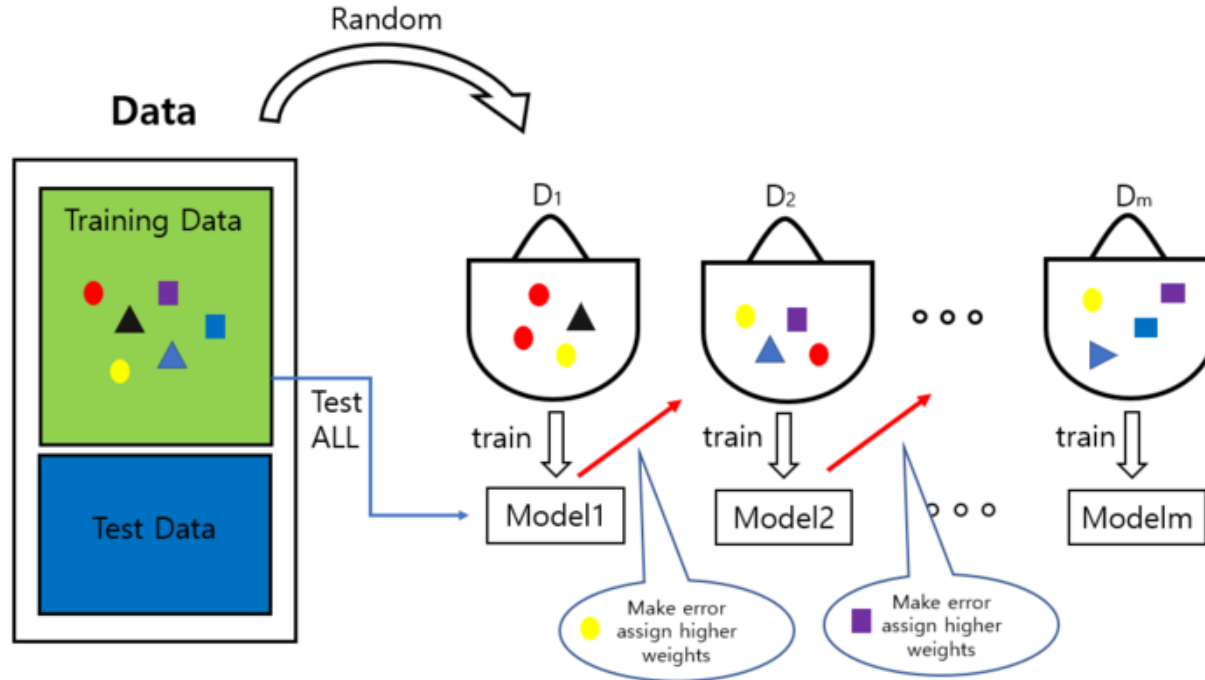
Adaptive Boosting의 약자.

이전 분류기가 잘 분류하지 못한 것들을

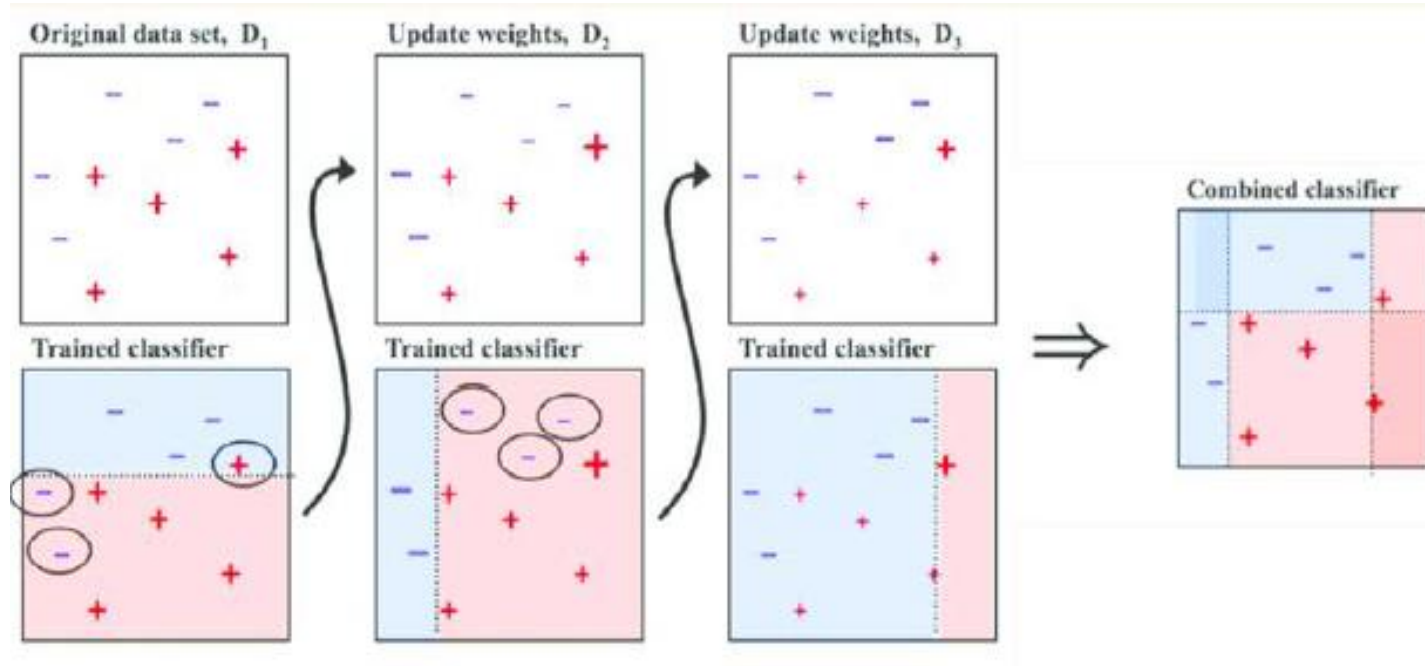
이어지는 약분류기들이 수정해 줄 수 있다는 점에서

다양한 상황에 적용할 수 있다.(Adaptive)

1. Ada Boosting



1. Ada Boosting



Ada Boosting 작동단계

현재 주어진 데이터셋에 대해 상대적으로 단순한 모델을 이용하여 학습



학습오류가 큰 개체의 선택확률을 증가, 오류가 작은 개체의 선택확률을 감소
-> 반복적으로 오류가 큰 데이터에 집중



일반적으로 분류모델의 경우는 Weighted Majority Vote,
회귀모델의 경우에는 Weighted Sum을 사용하여 최종 모델을 생성

Ada Boosting 특징

Bias를 줄이는데 초점을 맞춘 기법



Outlier에 취약

(오류가 크게 발생하는 Outlier에
반복적으로 집중하기 때문)

과적합(Overfitting)의 문제



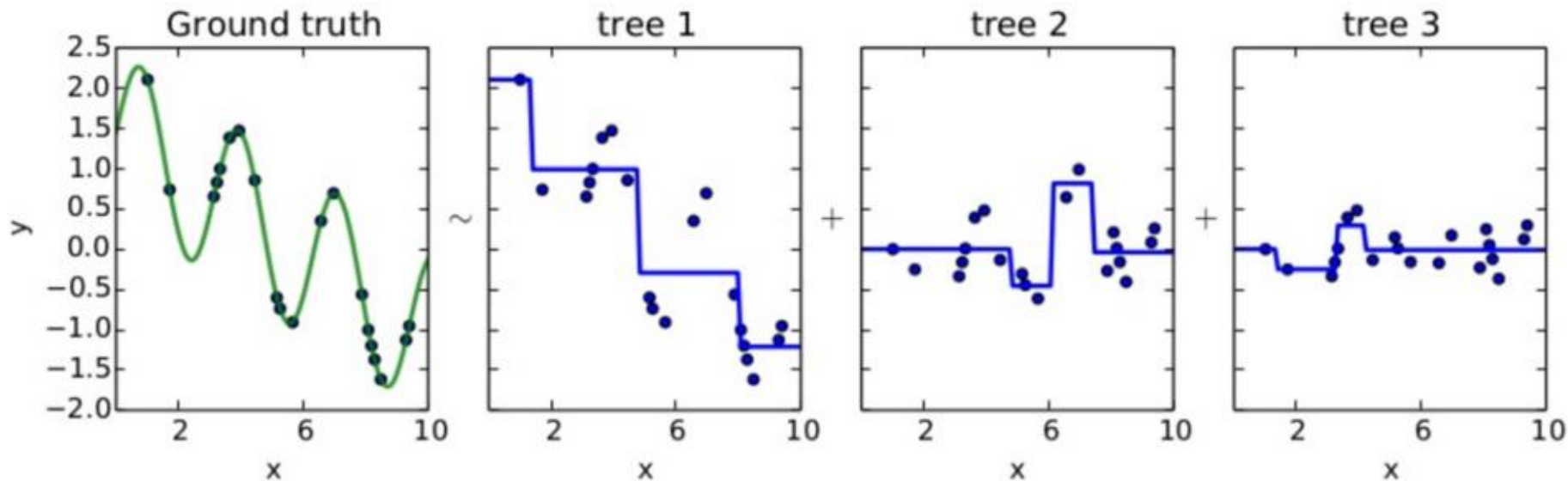
모델의 모수를 적절하게
조정하는 것이 중요!

2. Gradient Boosting

경사하강법을 사용해 손실함수(Loss function)를 최소화하는 가중치를 적용하여 Ada Boosting보다 성능을 개선한 Boosting 기법

- Outlier에 덜 민감하게 반응하므로 attribute(특성)의 scale을 조정하지 않아도 되고, continuous(연속적) attribute에서도 잘 작동
- Ada Boosting보다 학습시간 ↑ -> 적절한 변수 조정 필요
- Regression과 Classification 모두 적용 가능
- 고차원 데이터에서는 잘 작동하지 않을 수 있음

2. Gradient Boosting



분류기가 추가될 때마다 잔차 지속적으로 감소

-> Bias는 줄일 수 있으나 과적합(Overfitting) 발생 가능

-> sampling, penalizing 등의 regularization 기술을 이용하여 advanced된 모델 이용

3. XG Boosting

Extreme Gradient Boosting의 약자.

기존 Gradient Boosting의 **속도 문제를 해결**하기 위해

전산속도와 모델의 성능에 초점을 맞춘 기법

- 기존 Gradient Boosting에 비해 연산속도 빠르고 모델 성능 향상
- Gradient Boosting에 분산/병렬 처리가 추가된 알고리즘
- 과적합(Overfitting) 잘 일어나지 않음
- CPU와 같은 시스템 자원을 효율적으로 활용할 수 있는 알고리즘
- 최근 앙상블 모델 중에서 가장 우수한 알고리즘으로 평가받아 각종 대회에서 사용됨.