

인공지능 강의

머신러닝 활용 데이터 분석2

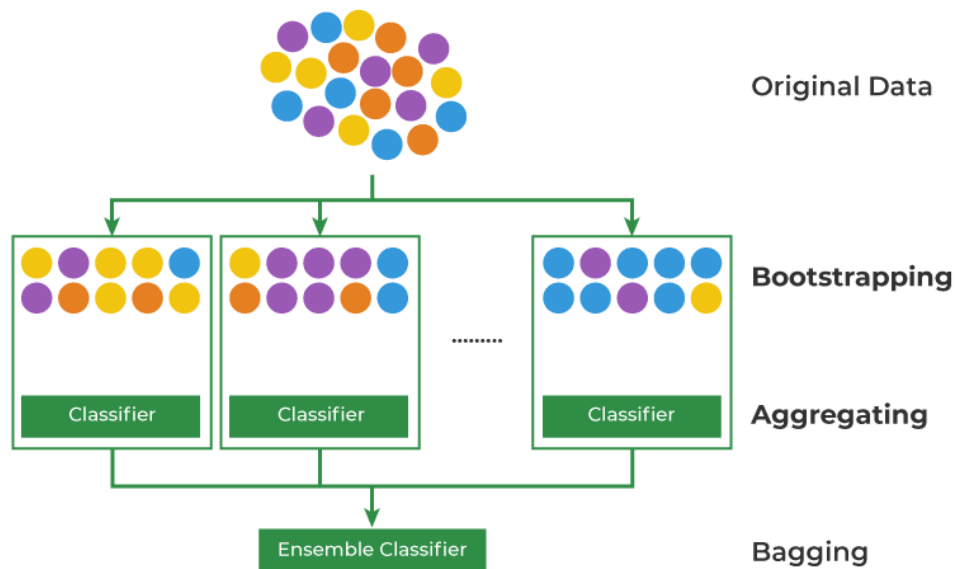


*모든 자료에 대한 권한은 메타코드에 있으며, 무단으로 자료를 복제 및 배포 등 유료목적으로 활용하시면 별도의 조치가 들어갈 수 있습니다.

Decision-Tree: Ensemble

‘Decision Tree를 합쳐서’
성능을 높이자

Bagging 알고리즘



Boosting 알고리즘



*모든 자료에 대한 권한은 메타코드에 있으며, 무단으로 자료를 복제 및 배포 등 유료목적으로 활용하시면 별도의 조치가 들어갈 수 있습니다.

Bootstrap 이란?

“Bootstrap은 통계학에서 사용되는 리샘플링(resampling) 기법”

*"Pull Yourself Up By Your
Bootstraps"*

BOOTSTRAPS



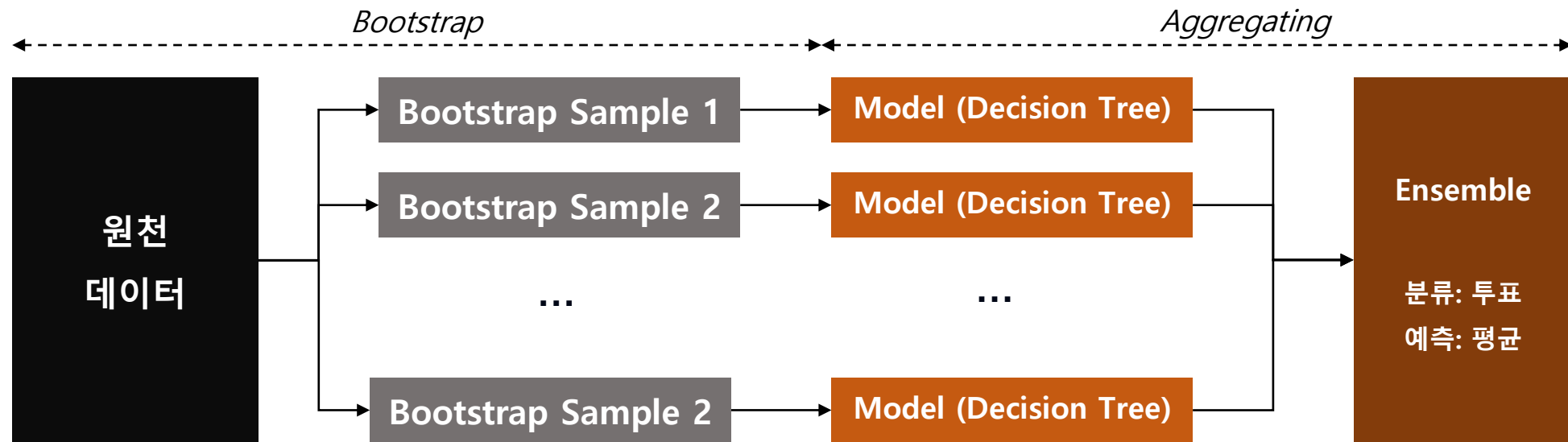
외부의 도움 없이 스스로의 힘으로 어려운
상황을 극복하고 성공하는 것을 의미



추가적인 외부 데이터 없이 원본 데이터셋
을 사용하여 새로운 데이터셋을 생성

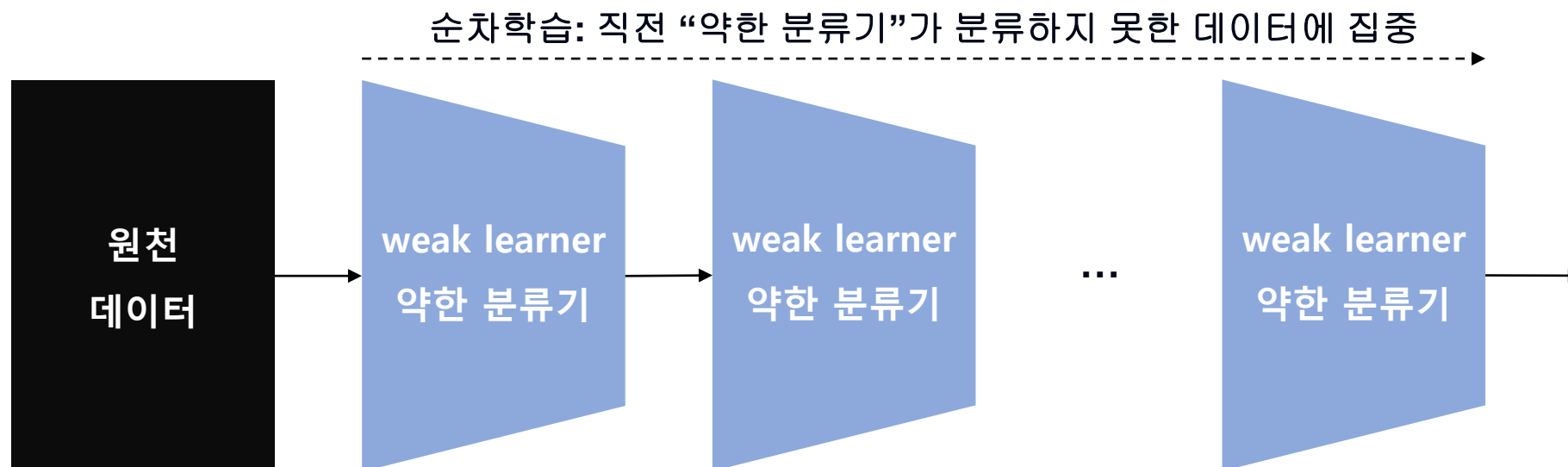
Bagging 이란?

“Bagging(Bootstrap Aggregating)은
Bootstrap 샘플링을 사용하여 여러 개의 모델을 학습하고
그 각각의 결과를 결합하여 최종 예측을 수행”



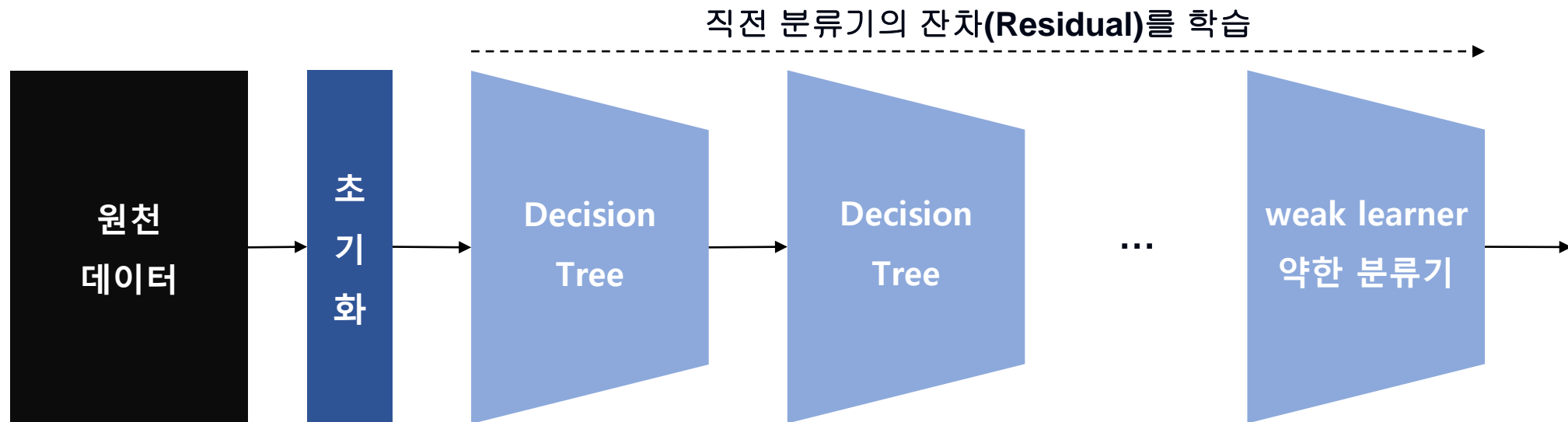
Boosting 이란?

“약한 학습기를 순차적으로 학습시키면서 각 학습기가 이전 학습기의 성능을 향상시키는 방식으로 앙상블”



Gradient Boosting 이란?

“직전 분류기가 분류하지 못한 데이터에 집중하여
다음 분류기는 그 데이터를 올바르게 분류하는 데 주력”



Gradient Boosting은 계속해서 이전 모델의 약점을 보완