

Part.04

Ensemble Learning

I Bagging

FASTCAMPUS
ONLINE

머신러닝과 데이터분석 A-Z

강사. 이경택

I Bagging

■ Ensemble Learning의 종류

- Bagging : 모델을 다양하게 만들기 위해 데이터를 재구성
- RandomForest : 모델을 다양하게 만들기 위해 데이터 뿐만 아니라, 변수도 재구성
- Boosting : 맞추기 어려운 데이터에 대해 좀더 가중치를 두어 학습하는 개념
Adaboost, Gradient boosting (Xgboost, LightGBM, Catboost)

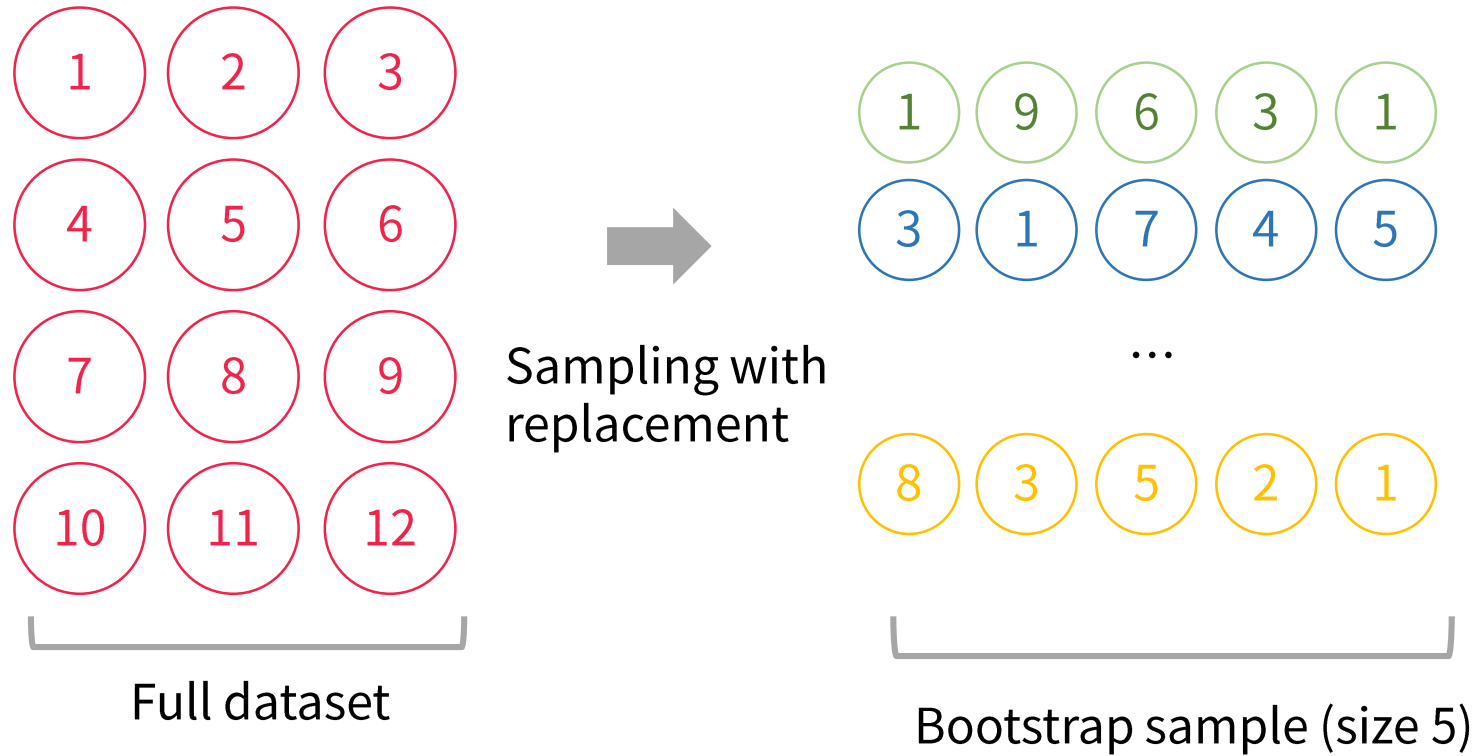
Tree기반의 단일 모델
(패키지 함수)

- Stacking : 모델의 output값을 새로운 독립변수로 사용

“Ensemble의 한 개념”

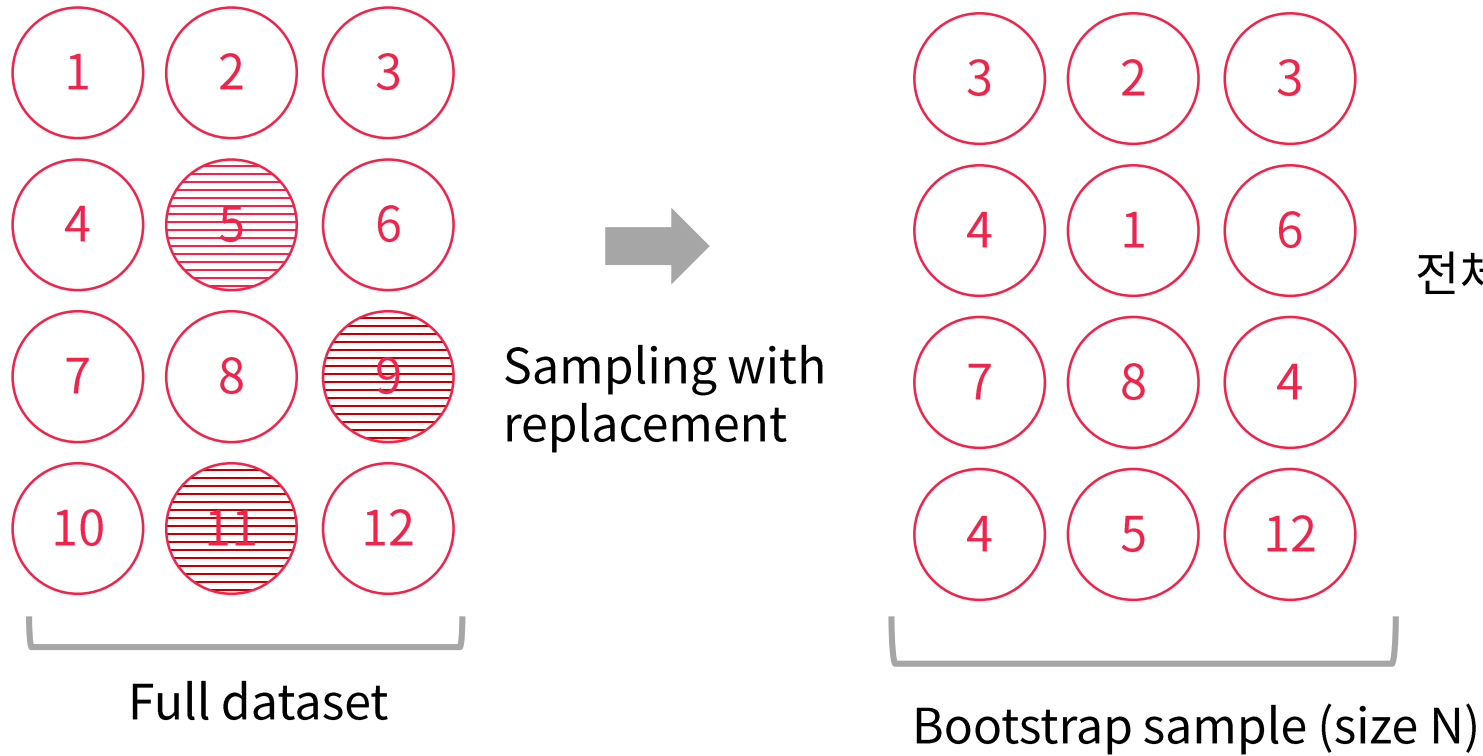
I Bagging

- Bagging (bootstrap aggregating)
 - Bootstrapping



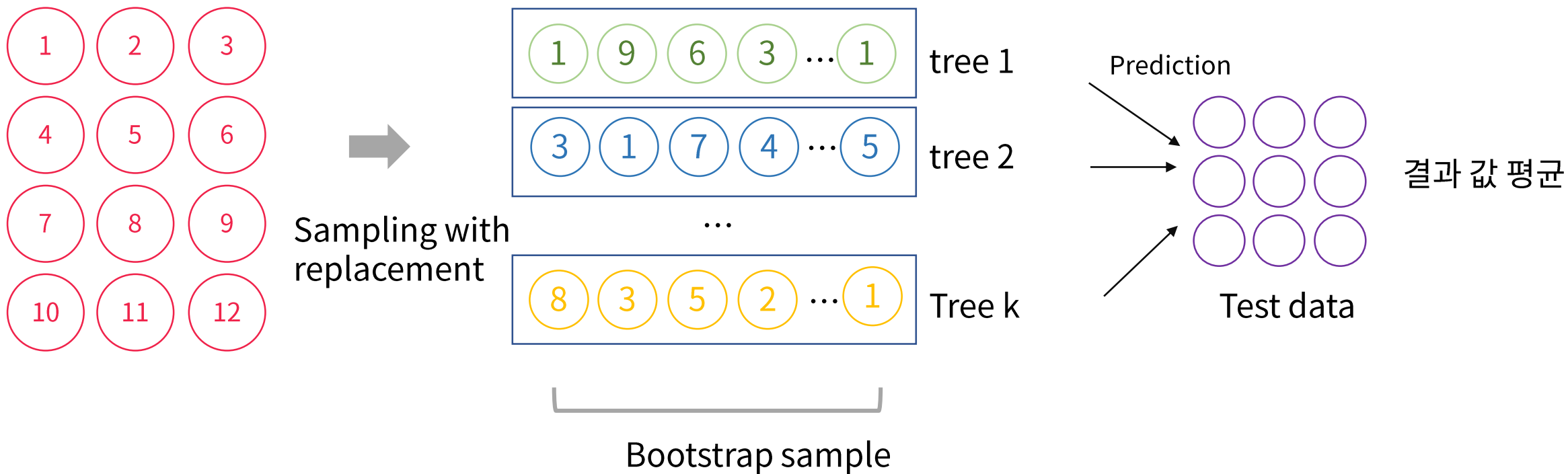
I Bagging

- Bagging (bootstrap aggregating)
 - Bootstrapping



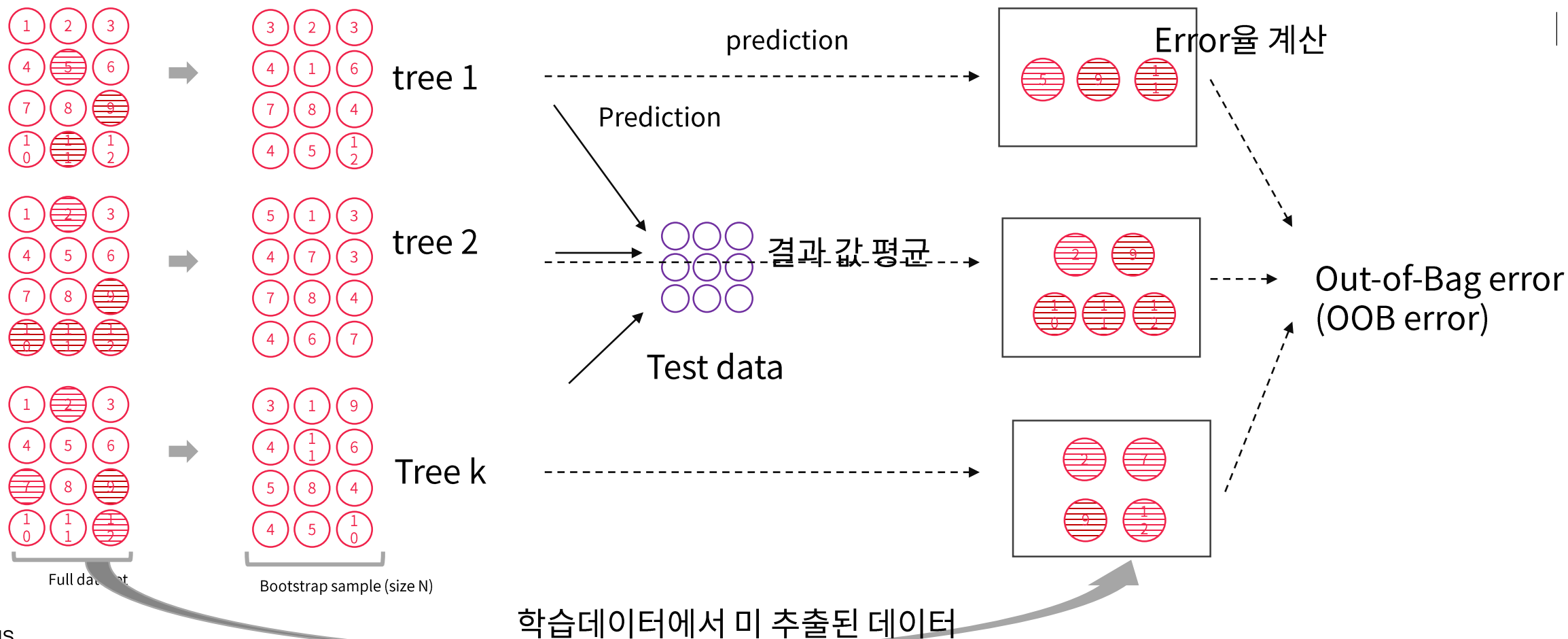
I Bagging

- Bagging (bootstrap aggregating)



I Bagging

■ Bagging (bootstrap aggregating)



I Bagging

- Bagging (bootstrap aggregating)

Tree vs Bagging

깊이 성장한 트리 : 분산 증가, 편향 감소

Bagging : 트리들의 편향 유지, 분산 감소 / 학습데이터의 noise에 강건해짐 / 모형해석의 어려움

Part.04

Ensemble Learning

I RandomForest

FASTCAMPUS
ONLINE

머신러닝과 데이터분석 A-Z

강사. 이경택