

Part.04

Ensemble Learning

I Ensemble의 Ensemble

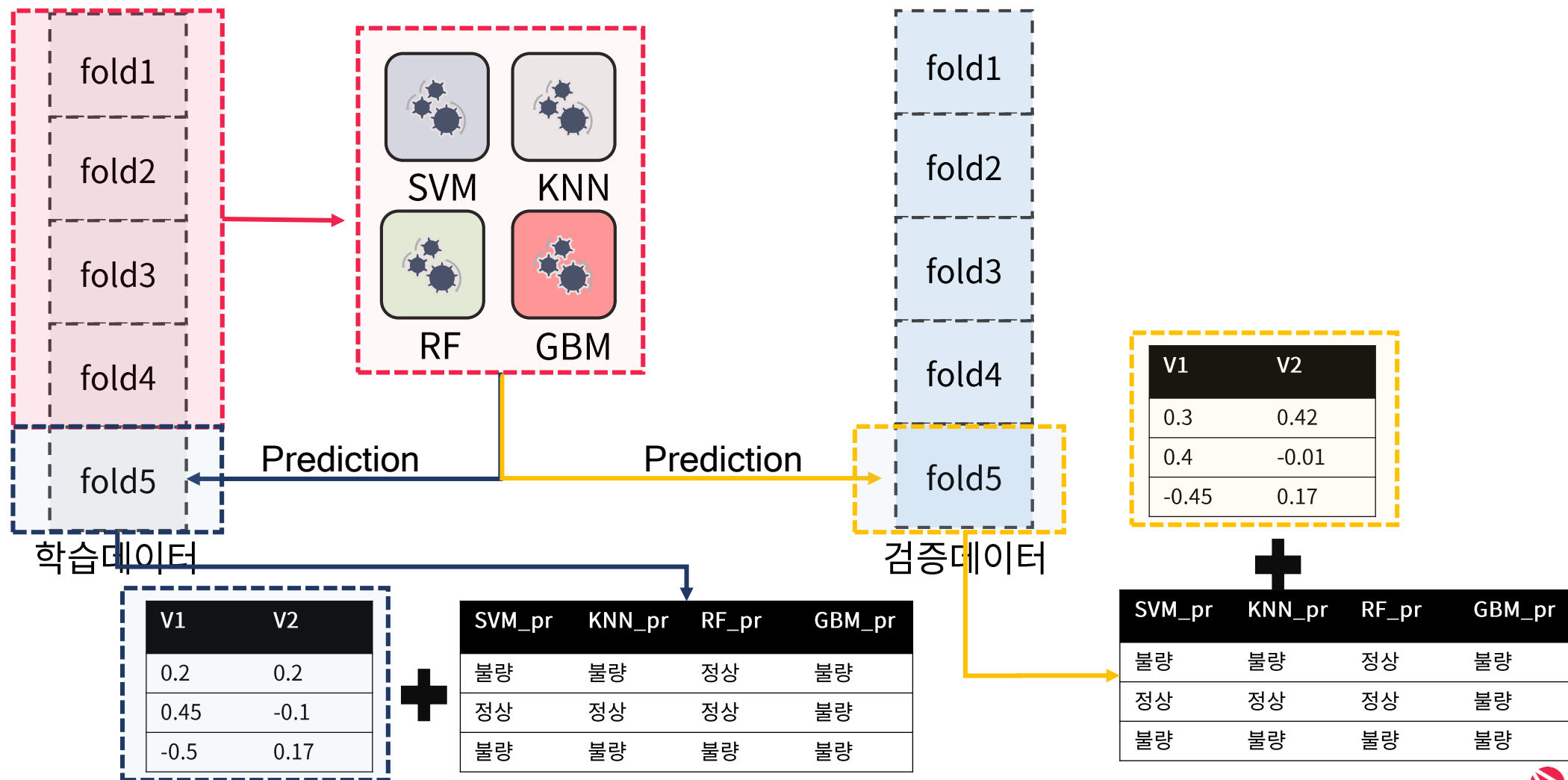
FASTCAMPUS
ONLINE

머신러닝과 데이터분석 A-Z

강사. 이경택

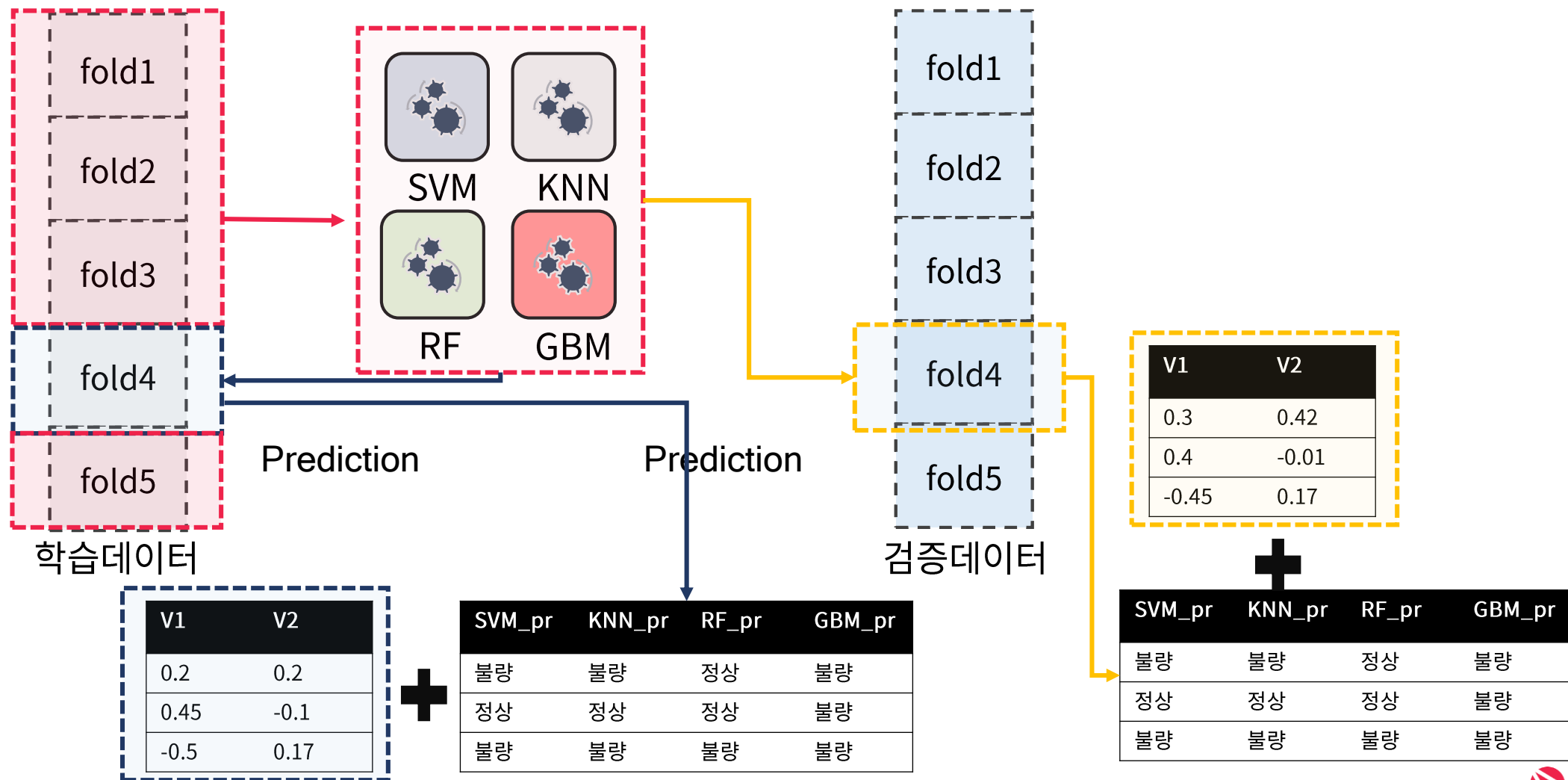
I Ensemble의 Ensemble

Stacking이란



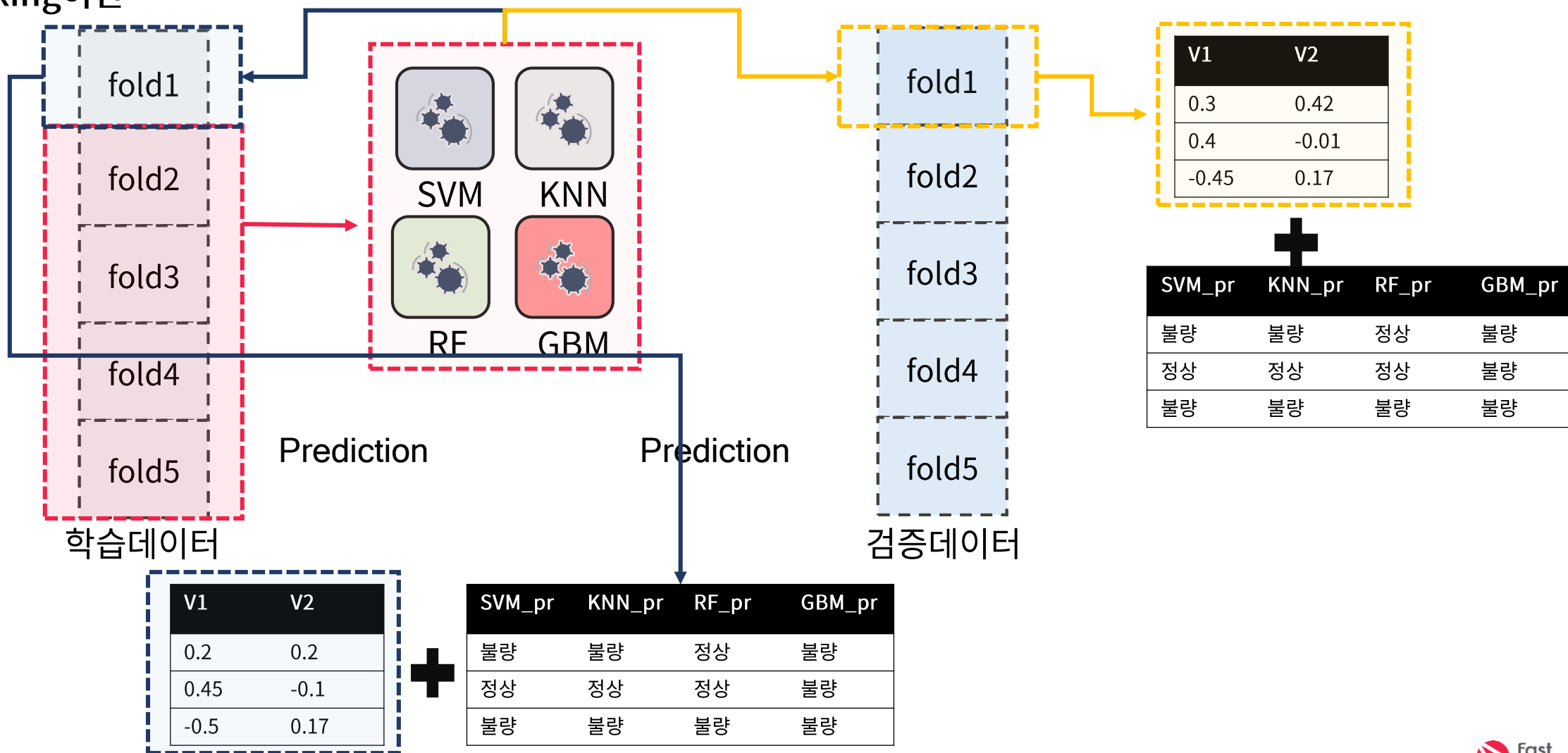
I Ensemble의 Ensemble

Stacking이란



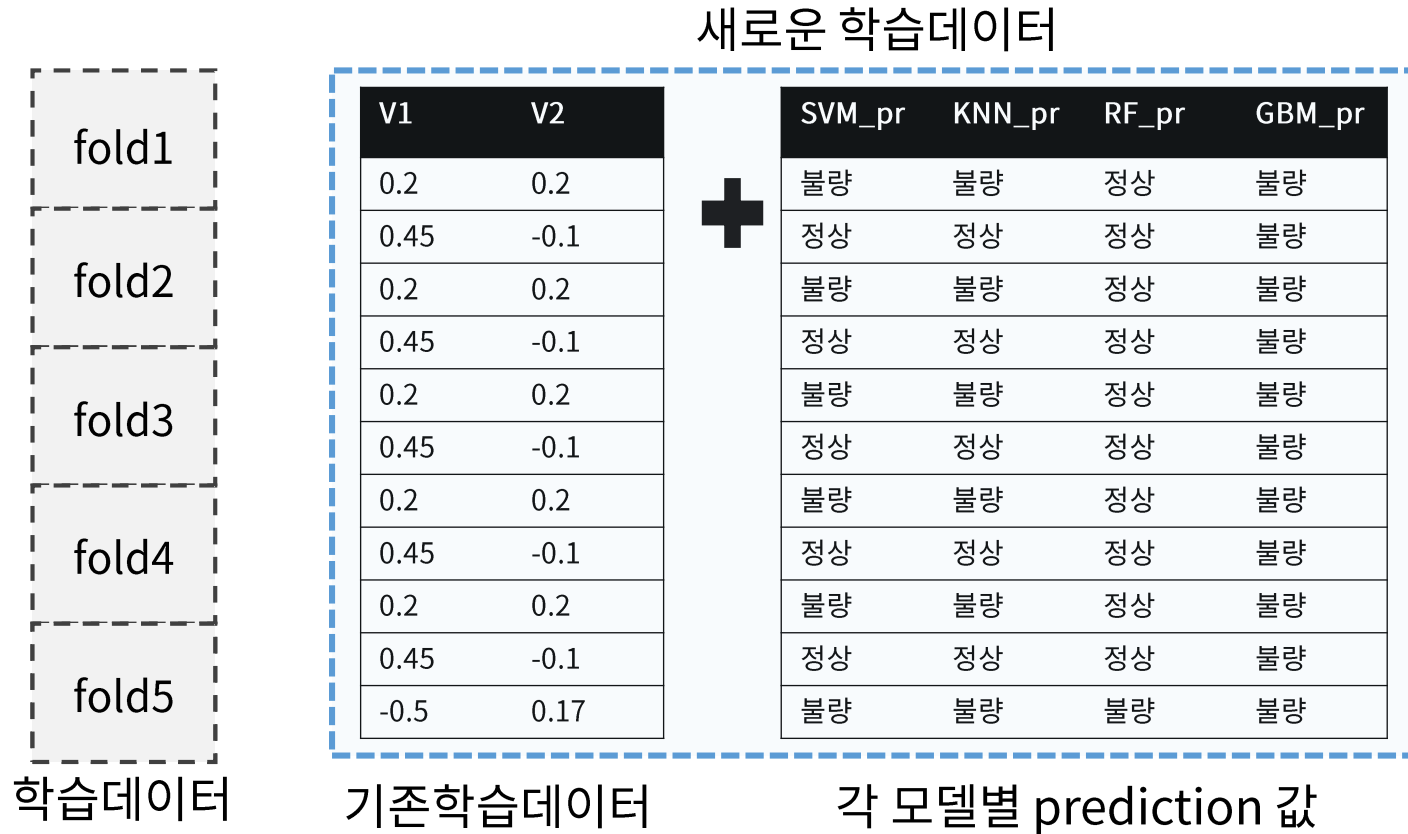
I Ensemble의 Ensemble

Stacking이란



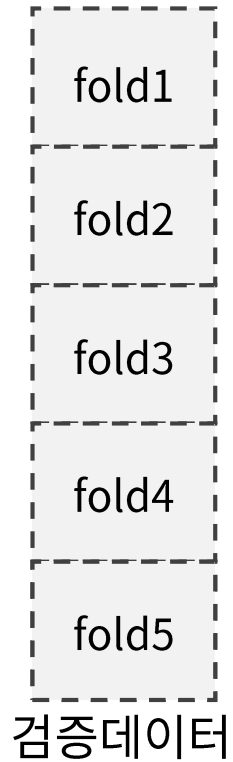
I Ensemble의 Ensemble

Stacking이란



I Ensemble의 Ensemble

Stacking이란



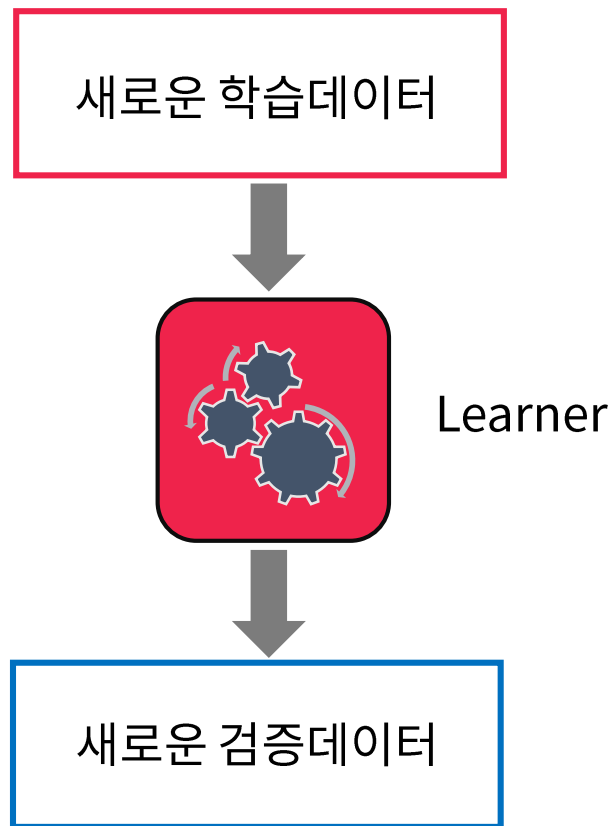
새로운 검증데이터

V1	V2	+	SVM_pr	KNN_pr	RF_pr	GBM_pr
0.2	0.2		불량	불량	정상	불량
0.4	-0.1		정상	정상	정상	불량
0.12	0.12		불량	불량	정상	불량
0.45	-0.1		정상	정상	불량	정상
0.2	0.12		불량	정상	정상	불량
0.35	-0.15		정상	정상	정상	정상
0.21	0.21		불량	불량	정상	불량
0.45	-0.1		정상	불량	불량	불량
0.02	0.2		불량	불량	정상	불량
0.45	-0.1		정상	정상	정상	정상
-0.5	0.17		불량	불량	불량	불량

기존검증데이터 각 모델별 prediction 값

I Ensemble의 Ensemble

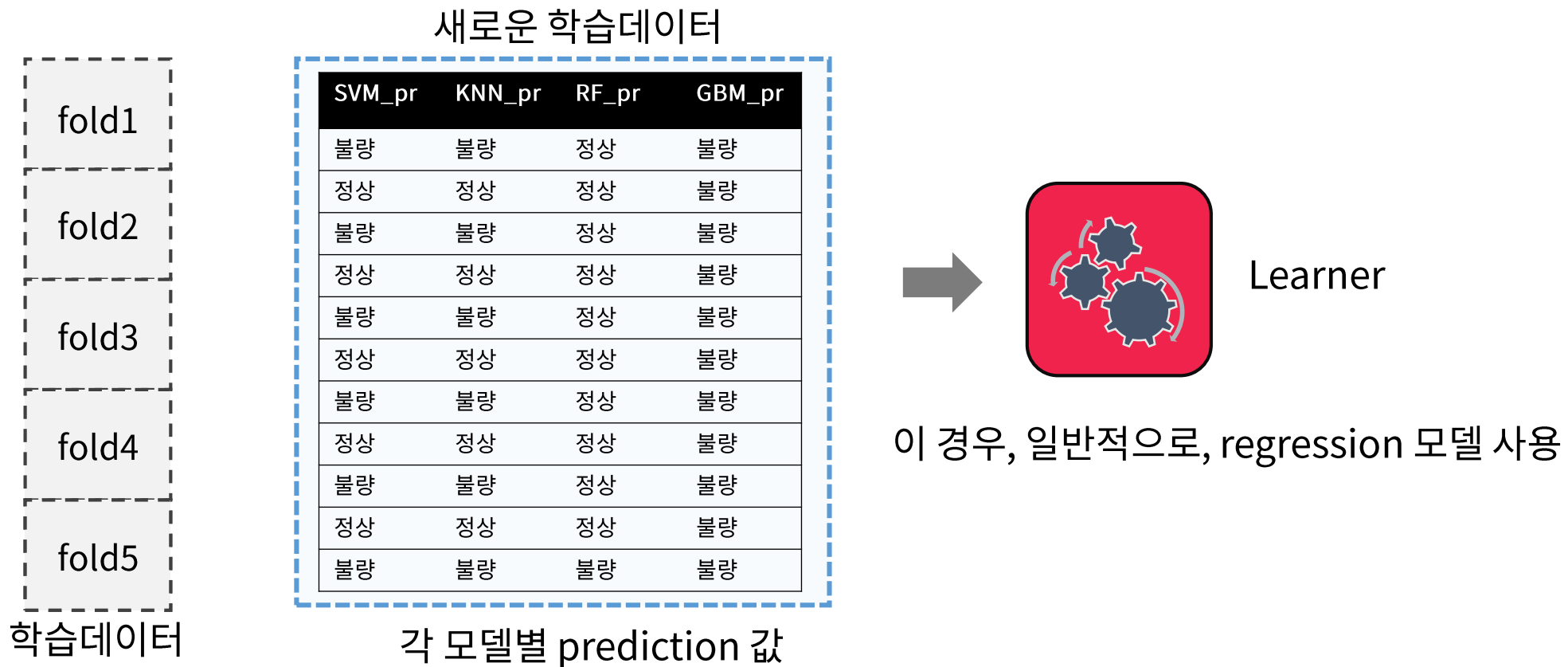
- Stacking이란



I Ensemble의 Ensemble

Stacking이란

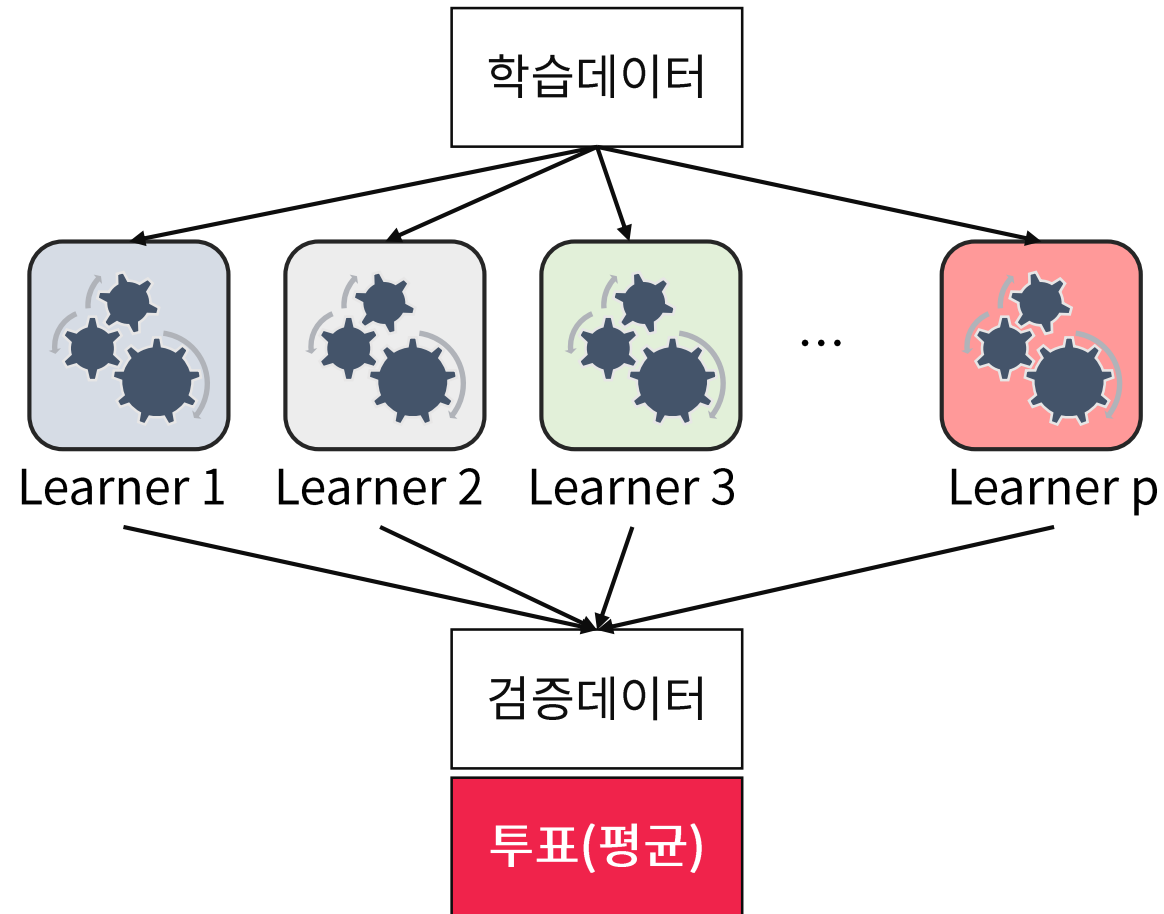
- 기존 feature를 쓰지 않고 각 모델 별 prediction만을 사용하기도함



I Ensemble의 Ensemble

■ Ensemble Learning

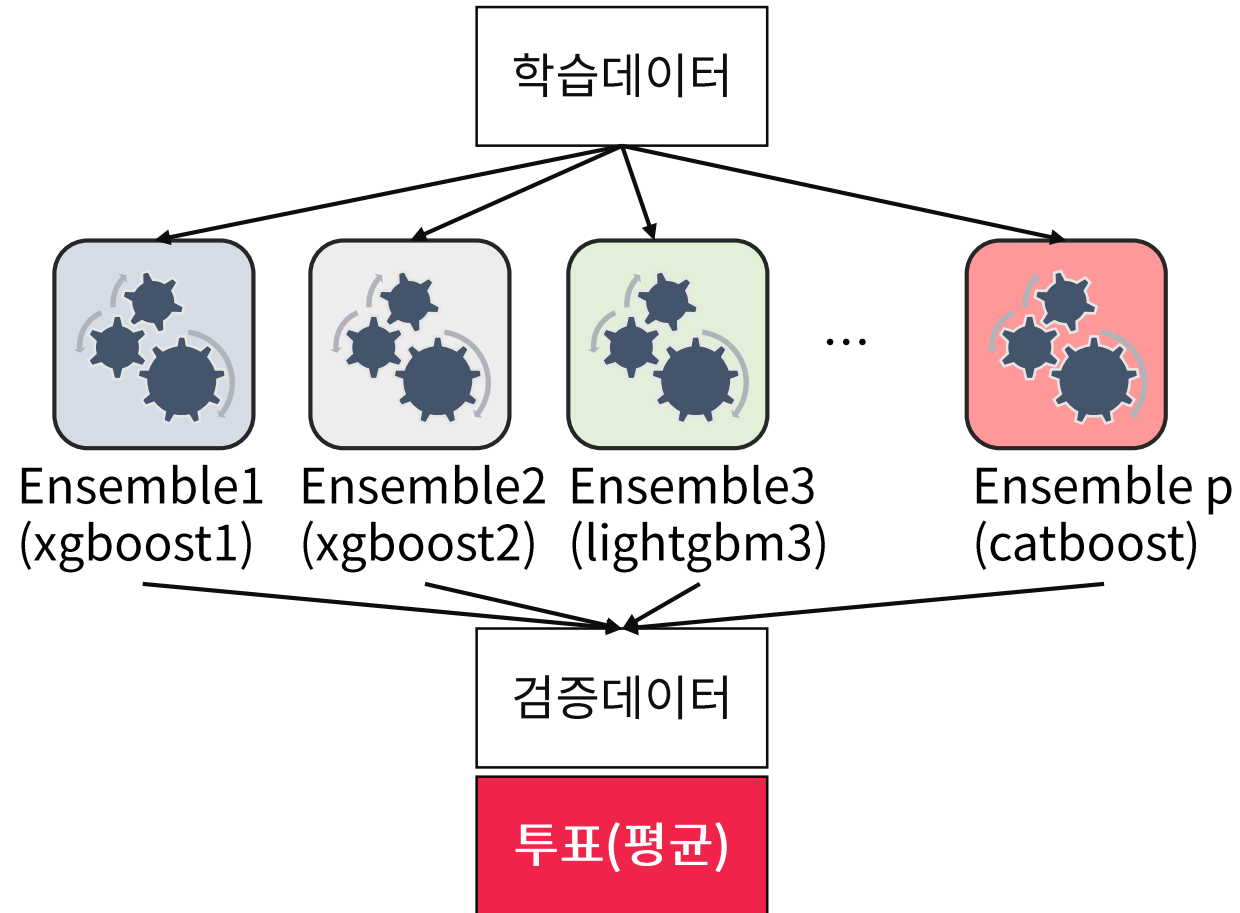
- 여러 개의 기본 모델을 활용하여 하나의 새로운 모델을 만들어내는 개념



I Ensemble의 Ensemble

■ Ensemble의 Ensemble

- Ensemble 모델을 단일 모델로 사용하자



I Ensemble의 Ensemble

■ Ensemble의 Ensemble

- Ensemble의 기본조건 – “다양한 모델”
- Boosting계열 알고리즘들은 hyper parameter에 민감한 경향이 있음 -> hyper parameter 다양화
- Bagging, RandomForest 컨셉과 같이 데이터 및 변수 random추출
- 일반적으로 단일 ensemble 모델에 비해 성능이 좋음

Part.04

Ensemble Learning

Ensemble의 기법 review

FASTCAMPUS
ONLINE

머신러닝과 데이터분석 A-Z

강사. 이경택