

앙상블 - 2인 이상의 합창이나 합주!

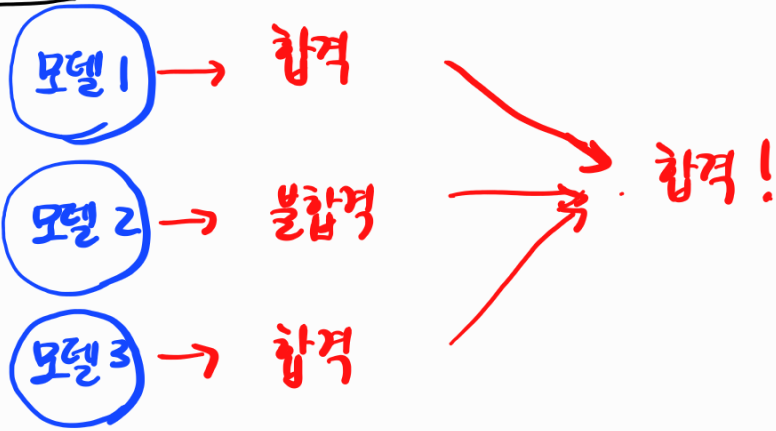
머신러닝에서 꼭

⇒ 평범한 모델을 여러개 합치면, 더 좋은 성능을 낼 수 있다.

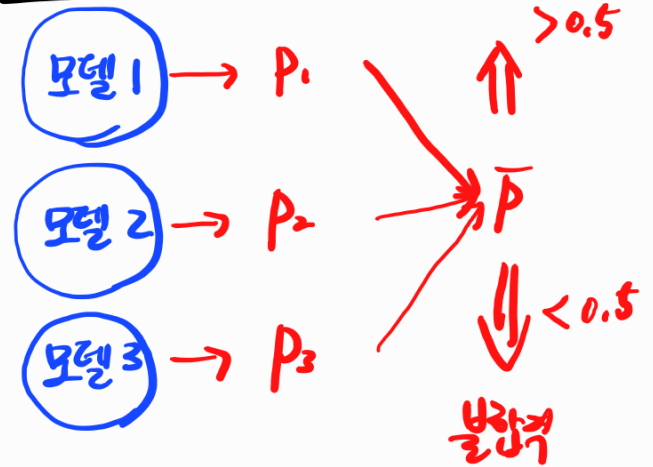
합치는 방법: 투표를 한다!

- Hard Voting (직접투표)
- Soft Voting (간접투표)

직접투표

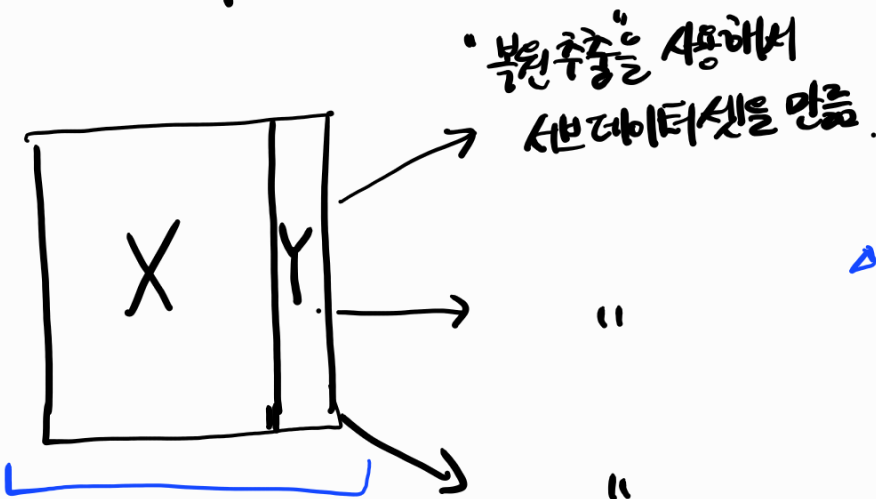


간접투표



\* 배깅 & ~~퍼이팅~~

(bootstrap aggregating)



✓  $X' | Y'$   $\xrightarrow{\text{fit}}$  모델 1

△  $X'' | Y''$   $\xrightarrow{\text{fit}}$  모델 2

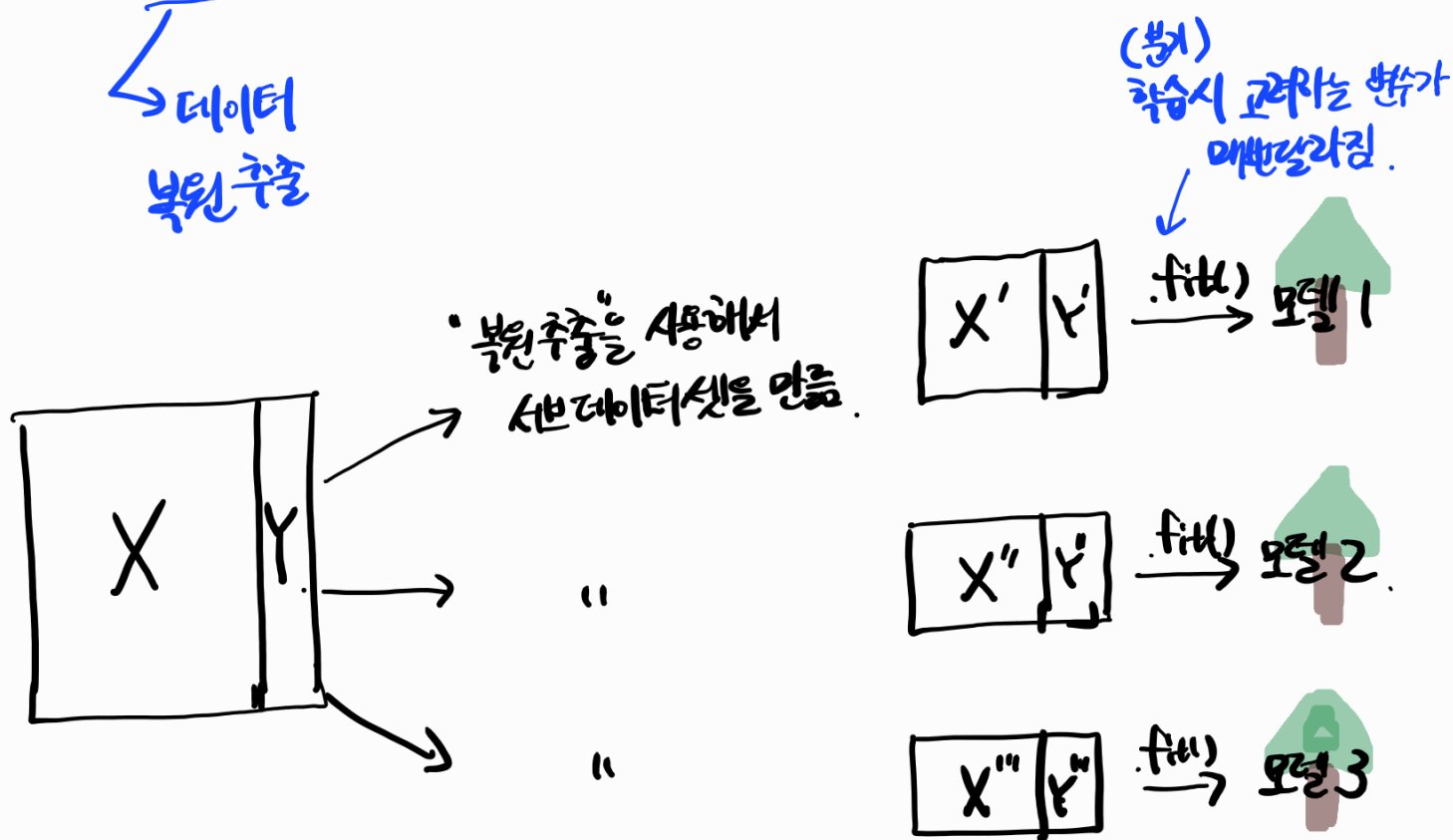
$X''' | Y'''$   $\xrightarrow{\text{fit}}$  모델 3

Bagging을 구현하는 사이킷런 패키지가 존재

# Random forests.

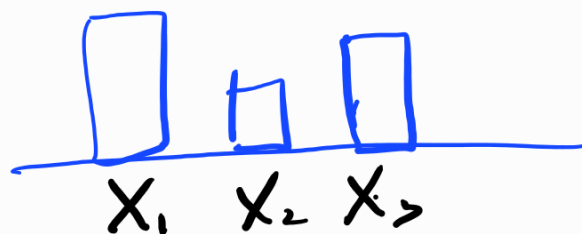
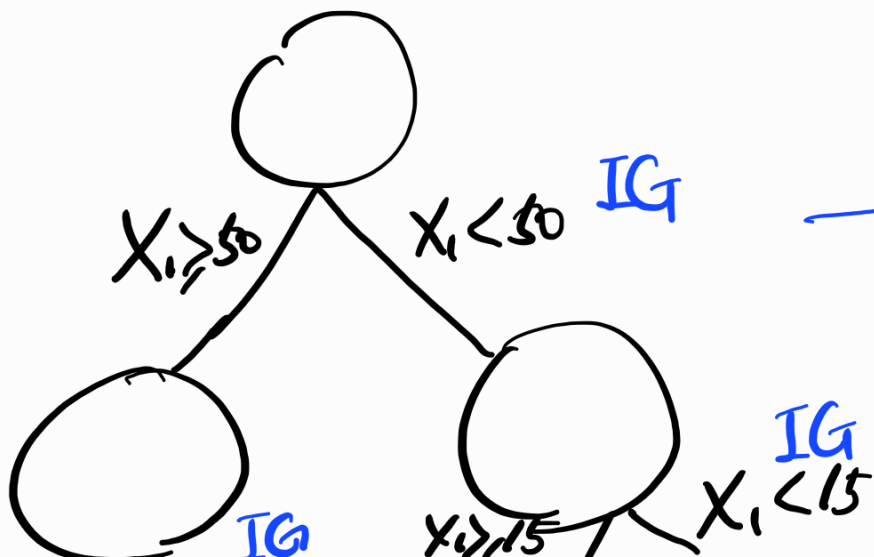
여러개의 tree 를 생성해서 하나의 앙상블 모델을 만듦.

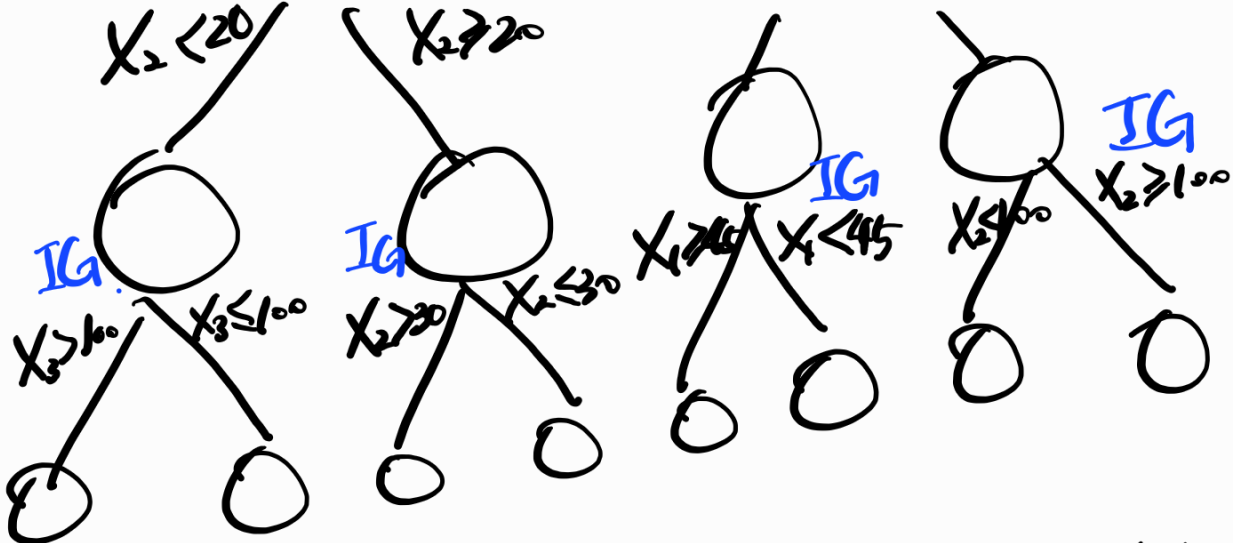
- Bagging 기법을 결정트리 모델을 생성해서 구현.



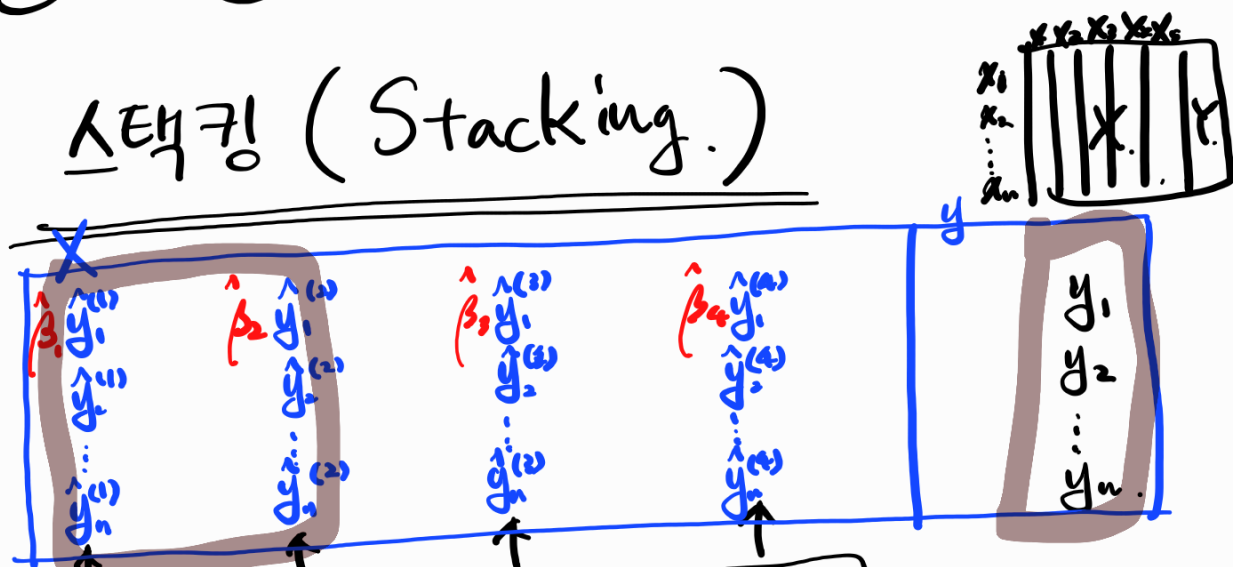
## 결정 트리 VS 랜덤 포레스트

$\Rightarrow$  각 분기 결정할 때, 변수를 모두 고려 (결정트리) VS 변수를 랜덤 선택 후 고려 (Random forests)



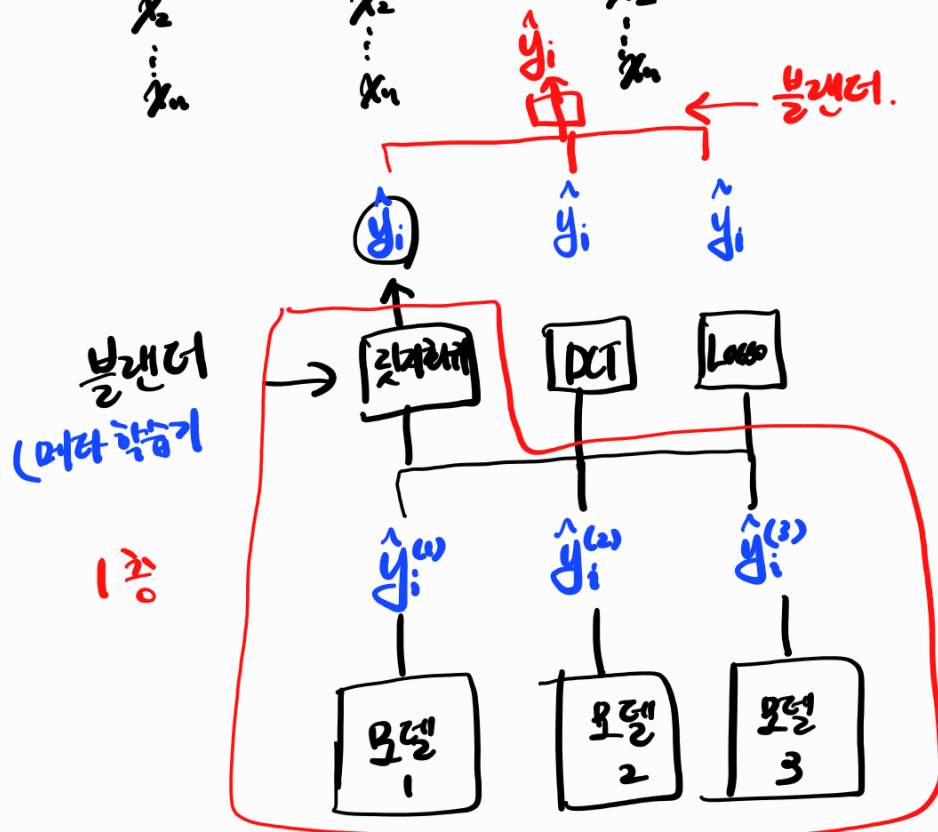


## 스택킹 (Stacking.)



학습에 필요한 데이터

예측할 때 입력 데이터.



1 2 3 NA 5 4

$\Rightarrow 3.$

