

앙상블

ensemble

앙상블 개념

- 주어진 자료로부터 여러 개의 예측 모형들을 만든 후
예측모형들을 조합하여 하나의 최종 예측모형을 만드는 방법
- 편향, 잡음 및 분산으로 인한 오류와 같은 문제를 회피하기 위해 사용됨
- 의사결정나무에 주로 사용 (과적합을 막기 위함)
- 배깅(bagging)과 부스팅(boosting), 랜덤 포레스트 (random forest)가 가장 대표적

앙상블 개념

- 스택킹 알고리즘 (앙상블과 유사하지만 예측 결과를 종합하는 것이 다른 점)
 - 여러 모델들을 활용해 각각의 예측 결과를 도출한 뒤
그 예측 결과를 다시 학습하여 최종 예측 결과를 만들어내는 방법

