13

배깅,부스팅 개념

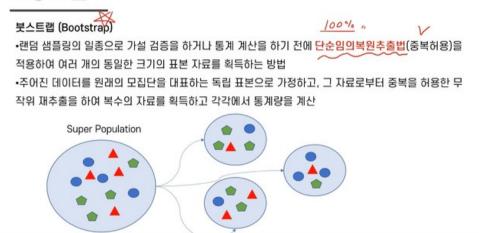
[배깅 개념 - 붓스트랩]

배깅을 이해하려면 붓스트랩의 개념을 잘 이해하셔야 합니다.

oob를 사용하면, 63%의 무작위 추출된 데이터를 train data로, 37%를 test데이터로 사용합니다.

00:07

배깅 개념



[배깅 개념]

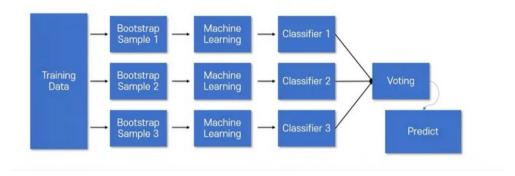
배깅에서는 여러개의 모델을가지고 투표하여 최종 예측데이터를 산출합니다.

● 배깅에서는 가지치기를 진행하지 않습니다.

02:28

배깅 개념

- •주어진 자료를 모집단으로 생각하여 주어진 자료에서 <mark>여러 개의 붓스트랩 자료를 생성</mark>하고 각 붓스트 랩 자료에 예측 모형을 만든 후 결합하여 최종 예측모형을 만드는 방법
- •분산을 줄이고 정확도를 개선하여 모델의 안정성을 크게 높여 과적합(overfitting)을 피하도록 함



[부스팅 개념]

부스팅은 배깅과는 다르게 직렬적으로 모델링이 수행됩니다.

- 가중치를 적용하여 데이터가 샘플링 될 확률을 바꿔줍니다.
- 하지만, 과적합이 될 위험이 있으므로 매개변수를 잘 설정해주어야 합니다.

06:52

부스팅 개념

- •예측력이 약한 모형들을 결합하여 강한 예측모형을 만드는 방법
- •붓스트랩을 병렬로 수행(각 모델을 독립적으로 구축)하는 배깅과 달리 순차방식으로 학습을 진행함

