

## 3 조 2 주차 발표자료\_앙상블

---

### 앙상블학습

여러 개의 분류기를 결합함으로써 보다 정확한 최종 예측을 도출하는 기법

다양한 분류기의 예측 결과를 결합 -> 단일 분류기보다 신뢰성이 높은 예측값을 얻음

---

### 보팅

여러 개의 분류기가 투표를 통해 최종 예측 결과를 결정하는 방식

서로 다른 알고리즘을 가진 분류기를 결합

하드보팅 : 예측 결과값 중 다수의 분류기가 결정한 예측값을 최종으로

소프트 보팅 : 분류기들의 레이블 값 결정 확률을 모두 더하고 평균내어 확률이 가장 높은 레이블 값을 최종 보팅 결과값으로 선정

➔ 일반적으로 소프트 보팅이 보팅 방법으로 적용

보팅 분류기 : 사이킷런의 보팅앙상블 VotingClassifier클래스

---

### 배깅

여러 개의 분류기가 투표를 통해 최종 예측 결과를 결정하는 방식

서로 같은 유형의 알고리즘 기반으로 하되 데이터 샘플이 각기 다른 분류기 사용

데이터 세트 간 중첩을 허용 (교차검증과 다른점)

---

### 부스팅

여러 개의 분류기가 순차적으로 학습을 수행하되 예측이 틀린 데이터에 대해서는 올바르게 예측할 수 있도록 가중치를 부여

➔ XGBoost, LGBM, ..

---

### 스태킹

스태킹은 여러가지 다른 모델의 예측 결과값을 다시 학습 데이터로 만들어서 다른 모델로 재학습시켜 결과를 예측

Dacon\_nuclear\_study\_1st\_3조

성능 수치를 조금이라도 높여야 할 경우 다시 사용

->많은 개별 모델 필요