

## 분류모델 요약

MNIST 데이터셋 - 0부터 9까지 숫자 이미지 데이터, 레이블 존재!

숫자를 레이블에 맞게 분류하도록 학습시키는게 목표이다.

문제 정의:

지도학습: 레이블(정답 존재)

분류: 0부터 9까지 숫자로 분류

배치 또는 온라인 학습: 둘 다 가능

훈련:

MNIST 데이터셋은 이미 훈련 셋과 테스트 셋이 6:1로 분류되어 있음

이진 분류기 예제 5-감지기:

1. 기존 이미지 레이블을 0으로 바꾸고 숫자 5를 가리키는 이미지는 레이블을 1로 바꿈
2. 확률적 경사 하강법: 매우 큰 데이터셋을 처리하는데 장점이 있음. 훈련 샘플을 독립적으로 처리.
3. 성능 측정: 기준을 정확도로 교차 검증을 사용, 오차 행렬을 통해 정밀도, 재현율로 성능 평가

다중 클래스 분류기:

- i) 일대다 방식: 5-감지기처럼 모든 숫자에 대해 감지기를 만들고 각기 다른 10개의 이진 분류기를 실행. 이후 각 분류기의 결정 점수 중에서 가장 높은 점수를 받은 클래스를 선택
- ii) 일대일 방식: 조합 가능한 모든 일대일 분류 방식을 진행해 가장 많은 결투를 이긴 숫자를 선택