

< Skewed Datasets - Error metrics for skewed datasets >

ex) Rare disease classification

Train classifier $f_{w,b}(\vec{x})$ $\begin{cases} y=1 & \text{if disease patient} \\ y=0 & \text{otherwise} \end{cases}$

\Rightarrow Find that you've got 1% error on the test set (99% correct)

\downarrow 잘 작동한 알고리즘인가?

if Only 0.5% of patients have the disease?

\Rightarrow 즉, $y=1$ 인 데이터셋이 전체 데이터셋 ($(y=0) + (y=1) = 100\%$)의 0.5%에 불과하다면?

항상 $y=0$ 으로 결정하는 알고리즘만으로도 (non-learning algorithm) 0.5%의 에러를 달성할 수 있다.

\therefore 1% error를 가지는 모델이 반드시 좋은 모델이 아닐 수도 있다.

\Rightarrow 이렇게 class별 데이터 수 차이가 많은 데이터는 "skewed data"라고 하며

skewed data인 경우 단순히 classification error가 아닌 다른 error metrics를 사용하여 모델의 성능을 평가

< Precision & Recall >

- 이러한 rare한 모델 성능을 평가할 수 있는 metric: precision / Recall

$y=1$: rare class we want to detect

* Confusion Matrix

		Actual Class	
		1	0
Predicted class	1	True Positive	False Positive
	0	False Negative	True Negative

TP (True Positive)

: 모델이 실제 True인 장악을 True라고 예측

FP (False Positive)

: 모델이 실제 False인 장악을 True라고 예측

FN (False Negative)

: 모델이 실제 True인 장악을 False라고 예측

TN (True Negative)

: 모델이 실제 False인 장악을 False라고 예측

* Precision (정밀도)

: 모델이 Positive라고 예측한 것 중 실제로 Positive인 비율

$$= \frac{TP}{TP + FP} = \frac{TP}{\text{total predicted positive}}$$

* Recall (재현율)

: 실제 Positive인 것 중 모델이 Positive라고 예측한 비율

$$= \frac{TP}{TP + FN} = \frac{TP}{\text{total actual positive}}$$