

Confusion Matrix

Confusion Matrix (오차행렬)

Model \Rightarrow MSE \Rightarrow R2 ✓
 \Rightarrow Acc (classification)
 True. false.
 1 0

Sample. 1000
 100 : True
 900 : False

1000개
 3개 잘못
 Acc 99.7%

Loss function
 Matrix

		실제 정답 (Ground Truth)	
		True 1	False
분류 결과	True 1	True Positive	False Positive
	False 2	False Negative	True Negative

Model이 잘못

실제 상황 (ground truth)	예측 결과 (predict result)	
	Positive	Negative
Positive	TP (true positive) 옳은 검출	FN (false negative) 검출되어야 할 것이 검출되지 않았음
Negative	FP (false positive) 틀린 검출	TN (true negative) 검출되지 말아야 할 것이 검출되지 않았음

Precision(정밀도)

- 모델이 True라고 분류한 것 중에서 실제 True인 것의 비율
- 맑다고 예측했는데 실제로 맑은날인 경우
- 예측한 결과가 실제 결과와 얼마나 일치하는가

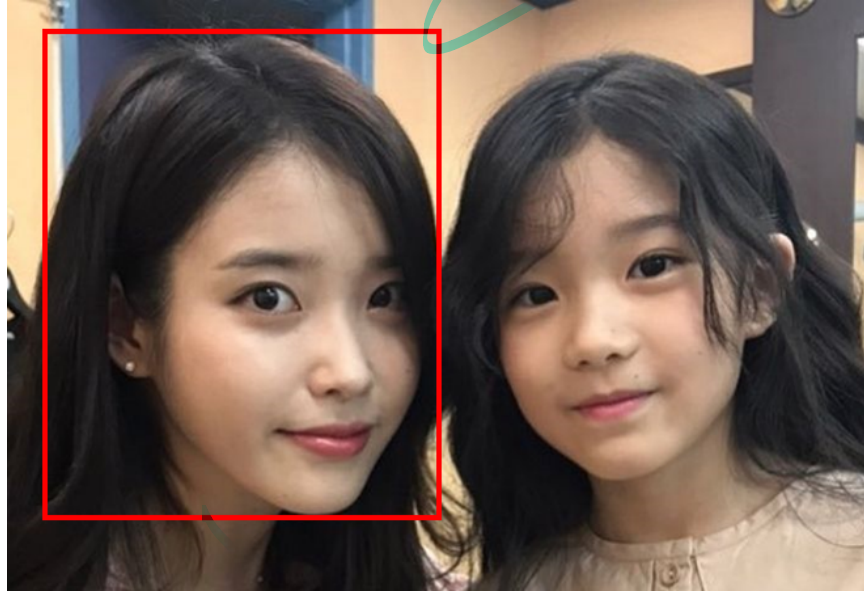
$$(precision) = \frac{TP}{TP + FP}$$

Recall(재현율)

- 실제 True인 것 중에서 모델이 True 라고 예측한 것의 비율
- 실제 맑은날 중에 모델이 맑다고 예측한 비율

$$(Recall) = \frac{TP}{TP + FN}$$

- Precision은 모델의 입장에서, 그리고 Recall은 실제 정답(data)의 입장에서 정답을 정답이라고 맞춘 경우를 바라보고 있습니다.
- 예측을 Face 로함 Precision = 100%
- 그러나 Recall = 50%



예) *의도: recall 중요. 환자가 제대로 판정받는지 확인*

A 병원 (잘된 분류 / 잘못된 분류)

실제값	예측값	
	암환자	일반환자
암환자	9	1
일반환자	30	60

B 병원 (잘된 분류 / 잘못된 분류)

실제값	예측값	
	암환자	일반환자
암환자	1	9
일반환자	20	70

recall

두개이레한

$$\text{precision} : \frac{9}{9+30} = \frac{9}{39}$$

$$\text{recall} : \frac{9}{9+1} = \frac{9}{10}$$

precision

recall

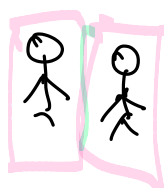
$$\text{pre} : \frac{1}{1+9} = \frac{1}{10}$$

$$\text{recall} : \frac{1}{1+20} = \frac{1}{21}$$

precision

구분화

2. 정확도 (accuracy)



- 모델이 입력된 데이터에 대해 얼마나 정확하게 예측하는지를 나타낸다.

$$\text{정확도} = \frac{\text{예측값결과와 실제값이 동일한 건수}}{\text{전체 데이터수}} = \frac{(TP+TN)}{(TP+TN+FN+FP)}$$

예)

$$\text{A병원의 정확도} = \frac{(9+60)}{(9+60+1+30)}$$

$$= 0.69$$

$$\text{B병원의 정확도} = \frac{(1+70)}{(1+70+9+20)}$$

$$= 0.71$$

정확도는 B병원이 높지만 암을 기준으로 하면 B가 더 안좋다.

3. 정밀도 (precision)

- 모델의 예측값이 얼마나 정확하게 예측됐는가를 나타내는 지표
- "예"라고 예측했을때의 정답률

$$\text{정밀도} = \frac{TP}{(TP+FP)}$$

예)

$$\text{A모델의 암환자 정밀도} = \frac{9}{(9+30)}$$

$$= 0.23$$

$$\text{B모델의 암환자 정밀도} = \frac{1}{(1+20)}$$

$$= 0.04$$

정확도는 B모델이 높지만 정밀도는 A모델이 높다.

4. 재현율 (recall)

- 실제값 중에서 모델이 검출한 실제값의 비율을 나타내는 지표
- Recall은 마땅히 검출해내야하는 물체들 중에서 제대로 검출된 것의 비율을 의미한다.
- 실제로 병이 있는 전체 중 참 긍정의 비율
- 실제 암환자들이 병원에 갔을때 암환자라고 예측될 확률,조기에 정확하게 발견해서 신속하게 처방하는 것이 올바른 모델

$$\text{재현율} = \frac{TP}{(TP+FN)}$$

예)

$$\text{A모델의 암환자 재현율} = \frac{9}{(9+1)}$$

$$= 0.9$$

$$\text{B모델의 암환자 재현율} = \frac{1}{(1+9)}$$

= 0.1

암환자 재현율을 기준으로 더 나은 모델은 A모델이다.

재현율 vs 정밀도

- 재현율과 정밀도는 사용하는 경우에 따라서 중요도가 다를수 있다.
- 재현율이 중요한 경우
 - 실제 Positive 인 데이터를 Negative로 잘못 판단하면 안되는 경우
 - **병 진단** : 실제 양성 인데 음성 으로 판단하면 병을 더 키울수 있다.
- 정밀도가 더 중요한 경우 (**precision 더 중요함**)
 - 실제 Negative 인 데이터를 Positive로 잘못판단하면 안되는 경우
 - 스팸 메일 : 실제 스팸메일이 아닌데(Negative) 스팸메일(Positive)로 판단하는 경우 메일을 받지 못할수 있다.

Confidence Threshold

Precision-Recall 곡선

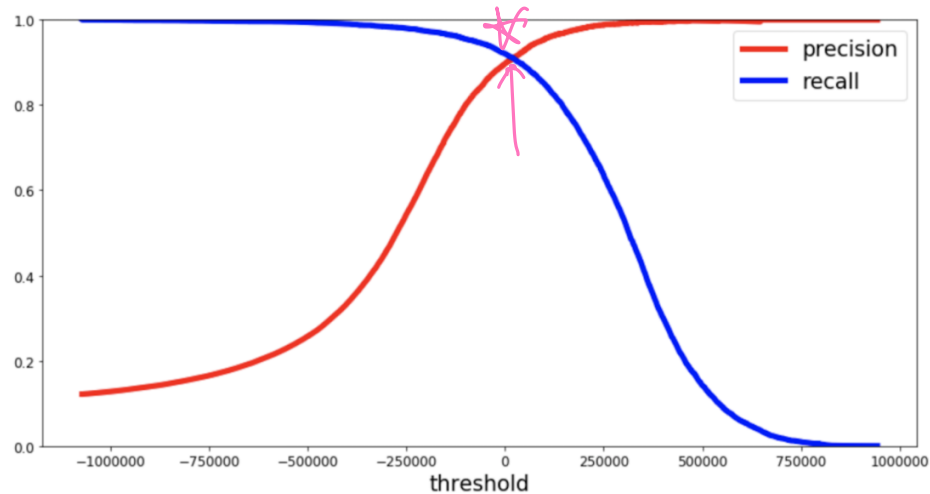
- PR 곡선은 confidence 레벨에 대한 **threshold** 값의 변화에 의한 물체 검출기의 성능을 평가하는 방법이다.
- confidence 레벨은 검출한 것에 대해 알고리즘이 얼마나 확신이 있는지를 알려주는 값
- 만약에 어떤 물체를 검출했는데 confidence 레벨이 0.999라면 굉장히 큰 확신을 가지고 검출한 것이다.
 - 알고리즘曰 "이 검출은 거의 99.9% 정확해. 난 그렇게 생각해."

참고

- confidence 레벨이 높다고 해서 무조건 검출이 정확한 것은 아니다.
- 알고리즘 스스로 그런 느낌 또는 확신을 갖고 있는 것
 - confidence 레벨이 낮으면 그만큼 검출 결과에 대해서 자신이 없는 것이다.
- 따라서 알고리즘의 사용자는 보통 confidence 레벨에 대해 threshold 값을 부여해서 특정 값 이상이 되어야 검출된 것으로 인정
 - threshold 값이 0.4라면 confidence 레벨로 0.1 또는 0.2를 갖고 있는 검출은 무시하는 것
- Confidence Threshold 값의 변화에 따라 정밀로-재현율은 변화한다.



- Confidence Threshold 값이 낮을 수록 더많은 예측 Bounding Box 를 만들게 된다.
 - (정밀도는 낮아지고 재현율은 높아짐)
 - Confidence Threshold 값이 높을 수록 신중하게 예측 Bounding Box 를 만들게 된다.
 - (정밀도는 높아지고 재현율은 낮아짐)
- Confidence Threshold 를 조정하면 정밀도 또는 재현율의 수치가 조정되고, 이는 서로 상보적이기 때문에 Trad-off 가 이루어 진다.



- Recall 값의 변화에 따른 Precision 값을 나타낸 곡선을 정밀도 재현율 곡선이라 한다.
- Precision 값의 평균을 AP 라고 하며, 면적값으로 계산 된다.

