



Chapter3. 평가

📅 날짜	@March 2, 2022
🏷️ 태그	파이썬 머신러닝 완벽가이드
📖 Property	p161~196 (pdf), p147~182(book)



머신러닝은 데이터 가공/변환, 모델 학습/예측, 평가의 프로세스로 구성된다.

여러 가지 방법으로 모델의 성능을 평가할 수 있고 분류 모델인지, 회귀 모델인지에 따라 성능 지표는 달라진다. 회귀는 실제값과 예측값의 오차 평균값에 기반한다. 분류는 여러가지 지표가 있다. 분류도 이진분류(binary), 다중분류(multi)에 따라 지표가 달라진다.

- 정확도 (Accuracy)
- 오차행렬 (Confusion Matrix)
- 정밀도 (Precision)
- 재현율 (Recall)
- F1 Score
- ROC AUC

▼ 1. 정확도(Accuracy)

실제 데이터에서 예측 데이터가 얼마나 **같은지**를 판단하는 지표

$$Accuracy = \frac{\text{예측결과가 동일한 데이터 건수}}{\text{전체 예측 데이터 건수}}$$

- 정확도를 통해 직관적으로 모델을 평가할 수 있지만, 이진 분류일 경우에 데이터의 구성에 따라 ML 모델의 성능을 왜곡할 수 있기 때문에 정확도만 가지고 평가하지 않는다. 정확도는 불균형한 label value 분포에서 ML 성능을 판단할 경우, 적합한 평가 지표가 아니다.

ex) 100개 중 90개가 0, 10개가 1일 경우 무조건 0으로 반환하는 모델의 정확도 : 90%

- 정확도가 가지는 한계점을 극복하기 위한 해결방법 : 다른 분류 지표를 함께 활용한다.

▼ 2. 오차 행렬(Confusion Matrix)

분류 모델이 예측을 수행하며 얼마나 혼동하는지까지 보여주는 지표 (예측 오류값 + 오류 유형)

Error Metrics

here: "setosa" = "positive"

actual class	setosa	TP 47	FN 3
	versicolor	FP 2	TN 48

setosa versicolor

predicted class

[Linear SVM on sepal/petal lengths]

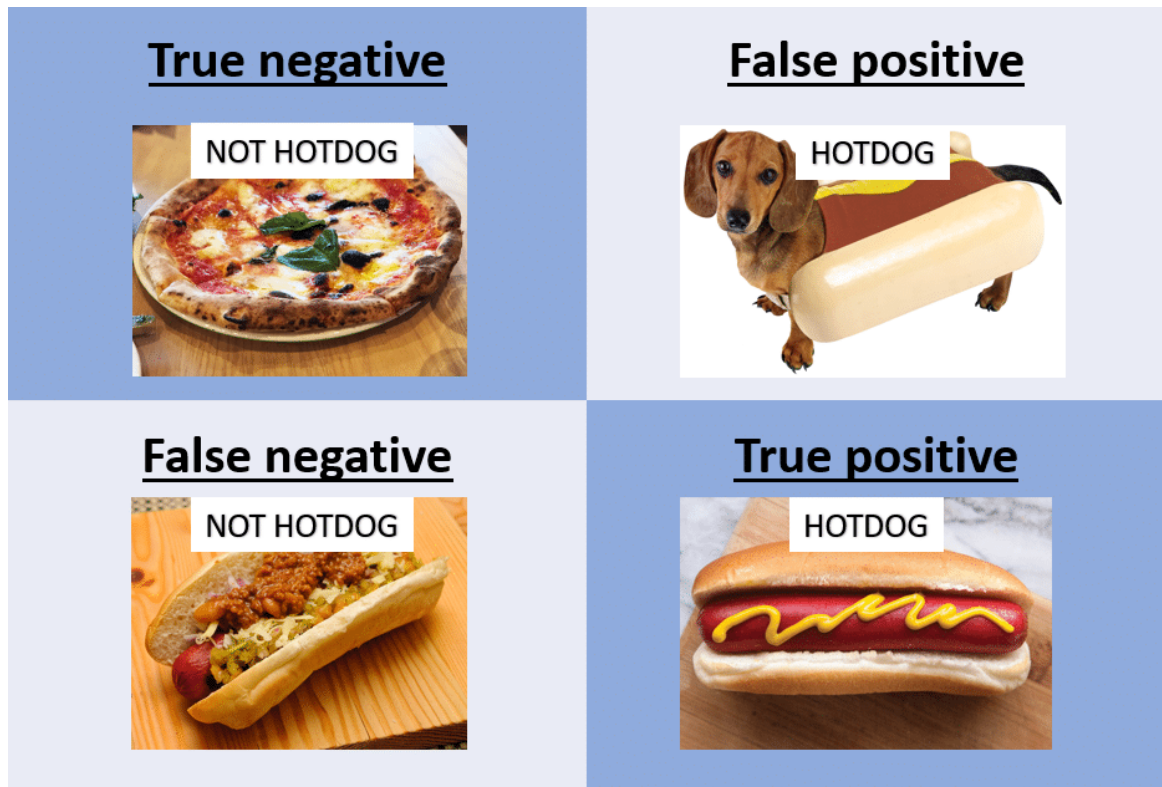
"micro" and "macro"
averaging for multi-class

$$Accuracy = \frac{TP + TN}{FP + FN + TP + TN} = 1 - Error$$

$$False Positive Rate = \frac{FP}{N}$$

$$True Positive Rate = \frac{TP}{P} \quad (Recall)$$

$$Precision = \frac{TP}{TP + FP}$$



- x축 : 예측된 클래스 값 기준 , y축 : 실제 클래스 값 기준
- 데이터가 불균형한 binary classification task에서는 positive의 데이터 건수가 작기 때문에 negative에 대한 정확도가 높아지는 현상이 발생할 수 있다. 99%가 negative, 1%가 positive라면 negative로 예측하는 경향이 더 강해져서 TN은 커지고 TP가 작아지는 현상이 발생할 수 있다. 이러한 사유로, positive에 대한 정확도를 판단하지 못한 채, negative에 대한 예측도가 높게 나타나는 경우가 있다.

▼ 3. 정밀도와 재현율

positive data set의 예측 성능에 초점을 더 맞춘 평가 지표

참고 : <https://jays-lab.tistory.com/31>

- 정밀도 : 예측을 Positive로 한 대상 중 예측과 실제값이 모두 positive로 일치한 데이터의 비율

$$\text{정밀도} = \frac{TP}{FP + TP}$$

- 재현율 : 실제값이 Positive인 대상 중 예측과 실제값이 모두 positive로 일치한 데이터의 비율로, 민감도 또는 TPR(True Positive Rate) 라고도 불림

$$\text{재현율} = \frac{TP}{FN + TP}$$

모델의 특성에 따라 어떤 것이 더 meaningful metric 인지는 달라질 수 있다. 재현율이 중요 지표인 경우는 실제 Positive 데이터를 Negative로 잘못 판단하게 되면 업무상 큰 영향이 발생하는 경우이다.

ex) 암 판단모델 : 오류의 담보는 생명... / 금융 사기 적발 모델 : 회사에 미치는 손해가 큼.

- 보통 재현율이 정밀도보다 상대적으로 중요한 업무가 많음. 정밀도가 더 중요한 경우는 실제 Negative data를 Positive로 잘못 판단하게 되면 큰 영향이 발생하는 경우.
ex) 스팸메일 : 스팸 메일을 일반 메일함에 넣는 경우
- 재현율의 초점 : FN 낮추기 / 정밀도의 초점 : FP 낮추기
- 서로 보완적인 지표지만, 가장 좋은 성능 평가는 두가지 지표 모두에서 높은 수치를 얻는 것이 보편적으로 좋은 모델이다. 어느 한 가지만 높고, 다른 건 낮은 경우는 바람직하지 않음.

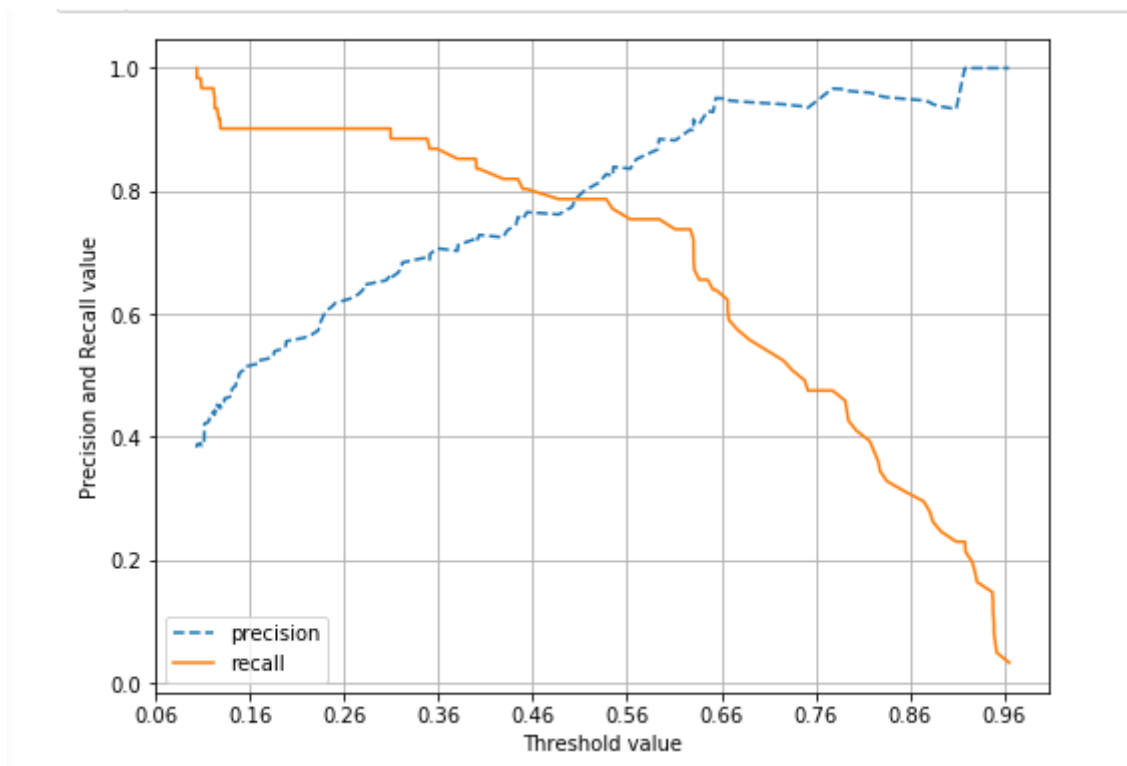
정밀도/재현율 Trade-Off

정밀도/재현율이 특별히 강조되어야 하는 경우 임계값(threshold)을 조정하여 수치를 높일 수 있다. 상호보완적인 지표이기 때문에 한 쪽을 높이거나 낮출 수 없다.

- 보통 sklearn에서는 prediction value가 큰 label 값으로 예측하게 되는데, 이진 분류에서는 threshold를 0.5(50%)로 정하는 것이 관례이다.
 - predict_proba() : output 값은 개별 클래스의 예측 확률 반환, 확률값이므로 0-1 사이의 값.
 - predict() : output 값은 예측 결과 반환 (labelled array). predict_proba()의 결과값 배열에서 **임계값보다 큰** 값이 들어있는 칼럼의 위치를 받아서 예측 클래스

를 결정하는 API

- 사이킷런의 `Binarizer()` 메서드를 사용하여 Trade-off 구현이 가능하다. `numpy ndarray`를 입력하여 입력된 임계값보다 작으면 0, 크면 1을 반환한다.
 - `binarizer(threshold=)`
- 실습(0.5 → 0.4) 에서 임계값을 낮추니 재현율 값이 올라가고 정밀도가 떨어졌다. `binarizer`의 `threshold`는 positive value 를 결정하는 확률의 기준이 된다. 예측을 더 쉽게(?) 하기 때문에 임계값을 낮출수록 True값이 많아지게 된다.
- 실습(0.4~0.6)에서 임계값 변화에 따른 평가 지표 값을 출력해보고 정확도와 재현율, 정밀도를 통해 적절한 임계값을 찾을 수 있었다. 또한, 임계값이 증가할수록 정밀도 값은 높아지나 재현율 값은 낮아짐을 알 수 있었다.



정밀도와 재현율의 맹점

positive 예측 임계값을 변경하면 정밀도와 재현율의 수치가 계속 변경되는데, 적당한 타협점을 잘 찾을 수 있어야한다. 단순히 성능 지표 수치를 높이기 위한 수단으로 사용해서는 안된다.

- 정밀도가 100%가 되는 방법 : 확실할 때만 positive, 나머지는 negative로 예측하기
- 재현율이 100%가 되는 방법 : 모든 환자를 positive로 예측하기

▼ 4. F1 Score

정밀도와 재현율을 결합한 지표로, 정밀도와 재현율이 균형을 이룰 때 상대적으로 높은 값을 가진다.

$$F1 = \frac{2}{\frac{1}{recall} + \frac{1}{precision}} = 2 * \frac{precision * recall}{precision + recall}$$

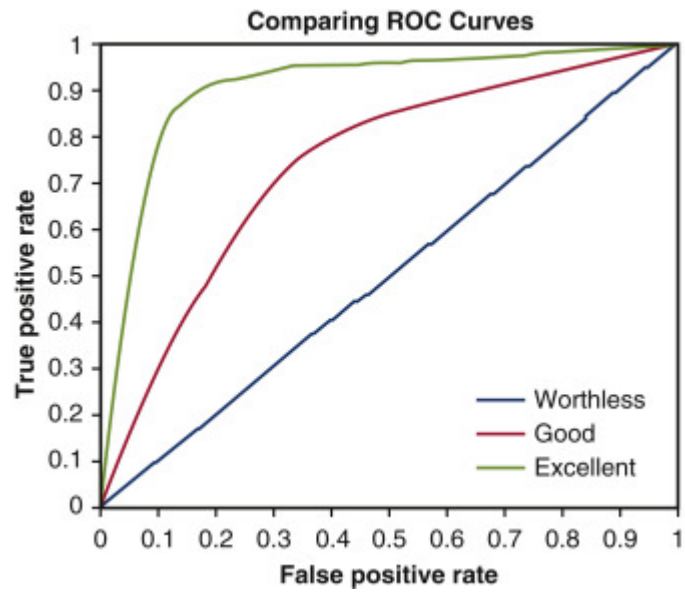
- A 모델의 차이가 크고, B 모델의 차이가 근소할 때 B 모델이 우수한 F1 score를 가진다.
- sklearn의 f1_score() 사용가능, 실습에서 확인

▼ 5. ROC Curve & AUC

ROC Curve와 AUC Score 는 이진 분류 예측 에서 중요하게 사용되는 지표다.

- ROC (Receiver Operation Characteristic) Curve == 수신자 판단 곡선
 - 원래는 2차세계대전 때 통신 장비 성능 평가를 위해 고안된 수치로, 의학 분야나 ML의 binary classification 에서 자주 사용된다.
 - **FPR(False Positive Rate) 이 변할 때 TPR(True Positive Rate)이 어떻게 변하는지 나타내는 곡선**이다. 각각 X축, Y축으로 잡고 그래프를 그리면 곡선 형태가 나타난다.
 - FPR(False Positive Rate) : $\frac{FP}{FP+TN}$. 1-TNR 과도 같다.
 - TPR (True Positive Rate) : 재현율. $\frac{TP}{FN+TP}$ 로 계산. 민감도라고도 불림.

- TNR (True Negative Rate) : 특이성. $\frac{TN}{FP+TN}$ 로 계산한다.



•